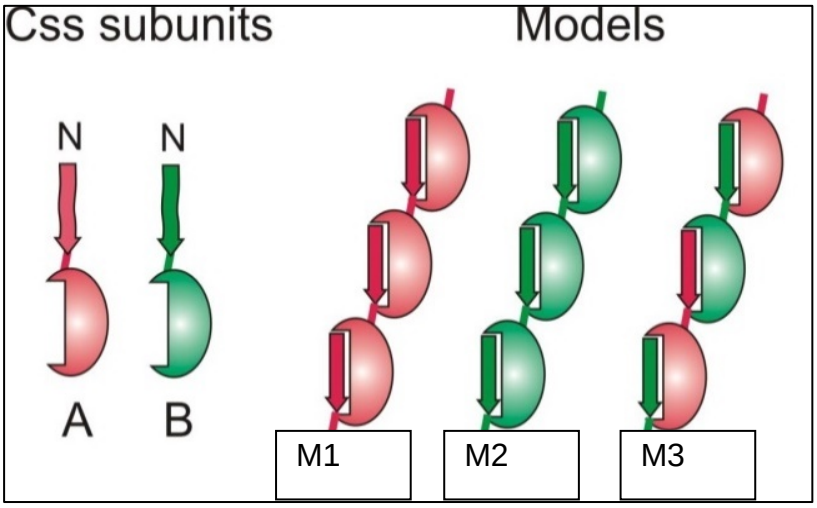
Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Molecular Sciences

Structure-Function study of CS6 fimbriae from ETEC

Taha Yousef Mahmoud Ghawanmeh



Independent Degree project in Biology- Master's thesis • 30 hec
Master's in Biotechnology
Molecular Sciences, 2018:28
Uppsala, 2018

Structure-Function study of CS6 fimbriae from ETEC

Taha Yousef Mahmoud Ghawanmeh

Supervisor: Anton Zavialov , Swedish University of Agricultural Sciences, Department of Molecular Sciences

Examiner: Jerry Ståhlberg, Swedish University of Agricultural Sciences, Department of Molecular Sciences

Credits: 30 hec
Level: BIA, Biology / Second cycle, A2E, Advanced
Course title: Independent Project/Degree Project in Biology - Master's thesis
Course code: EX0565
Programme/education: Master’s in Biotechnology

Place of publication: Uppsala
Year of publication: 2018
Title of series Molecular Sciences
Part number 2018:28
Online publication: https://stud.epsilon.slu.se

Keywords: ETEC, CS6, LT, ST, DSC , CD .

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Molecular Sciences

Abstract

Enterotoxigenic *Escherichia coli* (ETEC) are one of common pathogens causing diarrhoea in people worldwide. This bacterium is affecting mainly young children in developing countries and travellers to such destinations causing deaths and traveller's diarrhoea for them, respectively. The frequently isolated ETEC strains from patients have an immunogenic surface protein colonization factor (CS6). This surface antigen (CS6) is helping in adhesion to the host small intestinal epithelium enabling ETEC to produce the heat-labile (LT) and/ or heat-stable (ST) enterotoxin. The CS6 is expressed as typical bacterial operon that contains four genes *(cssA-BCD)* and transcribed once. The *cssA* and *cssB* encode for the structural subunits; *cssC* and *cssD* encode for chaperone and usher proteins, respectively. In order to make a preventive, treatment or vaccine strategy, more information is needed to be known about these genes expression and their surface assembly to form colonization factor and its role in attachment to the host intestinal epithelium.

The aim of this study is to express and purify the four possible structural subunits; CssA_dscA (AA), CssA_dscB (AB), CssB_dscB (BB) and CssB_dscA (BA) in order to determine their surface assembly and their stability functions. Future goals can be solving the 3D structures and designing candidate vaccines and/or treatments against ETEC. Methods used for this work include cloning of four possible structural subunits, protein expression tests, and protein purification by ion exchange and size exclusion chromatographic techniques and stability studies using CD (Circular Dichroism) spectroscopy.

In conclusion**,** the Donor strand complementation (DSC) theory was successful to predict candidate fimbrial structures that were purified and studied chemically and thermally in this work. The four subunits (AA, AB, BB and BA) were successfully cloned then protein expressed and purified using liquid chromatography techniques. Thermal and chemical studies showed that hetero-polymerized subunits, (AB and BA) were more stable against chemical and thermal denaturation compared to the homo-polymerized subunits (AA and BB). Subunit BA and AB were the most stable among the 4 subunits. These results favour the formation of hetero-polymerized subunits (AB and BA) more than the homo-polymerized ones. Further studies are needed to get better understanding of the mechanism of ETEC attachment to the host intestinal cells. These studies would need to get protein crystals then solving their structure in order to improve the chance of finding better treatment and successful vaccination method.

Keywords**:** ETEC, CS6, LT, ST, DSC**,** CD**.**

Dedication

To the Soul of my father “Yousef” for all of his Support, Courage, struggling and facing all the difficulties of the life, to give us a better life during all his life. May Allah bless his soul… forgive him of his sins… make his grave a garden and grant him the highest levels of paradise….. Amen.

To My struggling Mother for her Courage, Patience, Continuous and Incredible support throughout my life.

To all my sisters & brothers … for their support and encourage.

Finally and Always; to my lover, friend and soul mate, my beloved wife ENAS, and for my Small Angels and part of my heart, Tala and Qais.

Table of contents

List of tables **5**

List of figures **7**

Abbreviations **9**

1 Introduction **11**
1.1 Background 11
1.2 Role of Environment and Social factors : 12
1.3 Risk factors of ETEC infections: 12
1.4 Bacteriology: 13
1.4.1 Classification and serotypes: 13
1.4.2 Virulence factors and CFs – binding sites and their effect 13
1.5 Clinical Feature 14
1.6 Prevention and Vaccination 14
1.7 CS6 gene of ETEC: 15
1.8 Chaperone/usher pathway and Donor strand complementation theory 16
1.9 Chaperone structure and interaction with subunits 16

2 Materials and methods **19**
2.1 Expression Plasmids Construction: 19
2.1.1 Isolation of the (*AA, BB, AB* and *BB*) DNA fragments by PCR. 19
2.1.2 Agarose Gel Electrophoresis 20
2.1.3 DNA Purification from Agarose gel 20
2.1.4 Cloning and ligation in to pET101/ TOPO vector 21
2.1.5 Colony PCR 21
2.2 Plasmid Purification 22
2.3 Glycerol stock preparation 23
2.4 Transformation of plasmids into expression cells BL21-DE3 23
2.5 Screen of clones by test expression 23
2.6 Protein harvesting from periplasmic space 24
2.7 SDS-PAGE (15%) For protein expression analysis 24
2.7.1 Gel staining with Coomassie Blue: 25
2.8 Large culture and purification of protein 25
2.9 Dialysis and filtration of protein 26
2.10 Protein purification by liquid Chromatography 26

2.10.1 Anion-exchange chromatography (Source 30Q and MonoQ) 27
2.10.2 Mono Q chromatography (8 CV, 5MPa) 27
2.10.3 Cation exchange chromatography (Mono S) 27
2.10.4 Size exclusion chromatography (Gel filtration) 28
2.11 Protein Stability study by Circular Dichroism (CD) Spectroscopy. 28
2.11.1 Chemical denaturing using Guanidine Hydrochloride (GdnHCl) 28
2.11.2 Thermal stability of the studied protein subunits 29
2.11.3 Free energy change ΔG (kJ/mol) Calculation : 29

3 Results and Discussion 31
3.1 Construction of Expression Plasmid for the four complemented subunits 31
3.2 Test expression results on SDS-PAGE 32
3.3 Anion-exchange chromatography (Source 30 Q) results 33
3.4 Anion-exchange chromatography (Mono Q) results 35
3.5 Size exclusion chromatography (Gel filtration) 38
3.6 Results of Circular Dichroism (CD) Spectroscopy 40
3.6.1 Chemical denaturing of the CssA_dscA subunits. 40
3.6.2 Thermal stability. 41
3.6.3 Calculated free energy change, ΔG (kJ/mol): 42

4 Conclusion 45

5 References 47

6 Popular science summary 51

7 Acknowledgements 53

List of tables

Table 1. PCR reaction mixture ingredients used for AA and BB fragment isolation 20
Table 2. PCR program 20
Table 3. PCR reaction mixture of Colony PCR 22
Table 4. Colony PCR program 22
Table 5. SDS PAGE ingredients for separation and stacking gels 24
Table 6. List of Dialysis buffer used 26
Table 7. Reaction mix of GdnHCl and protein subunits by Circular Dichroism (CD) 29
Table 8. Summary of Chemical stability using (GdnHCl) needed for protein denaturation 41
Table 9. Summary of Melting Temperature needed for protein denaturation. 42
Table 10. Summary free energy change, ΔG (KJ/mol), needed for thermal denaturation at melting point (Tm) of each subunit. 43

List of figures

Figure 1. Versions of complemented subunits 17

Figure 2. The possible polymers of this study. 18

Figure 3. PCR results of 4 subunits isolation at 1 % Agarose gel: lane 1 represents the ladder (1000bp); lane 3, AA; lane 4, AB; lane 6, BB; and lane 7, BA. 31

Figure 4. Test expression results on SDS-PAGE: lane 1, AB cells control; lane 2, AB cell (none induced); lane 3, AB cell (induced); lane 4, AB cell sucrose fraction; lane 5, AB old cell (none induced); lane 6, AB cells (induced); lane 7, BA cells (none induced); lane 8, BA cells (induced); lane 9, BA cells sucrose fraction; lane 10, BA old cells (none induced); lane 11, BA old cells (induced); lane 12, AA cells (induced); lane 13, BB cells (induced) and lane 14, control AB cells. 32

*Figure 5*A. Source30Q chromatography results for BB subunits: absorbance at 280 nm, and peak is between fraction numbers 16 and 22. 33

Figure 5B. SDS–PAGE results from peak fraction of the Source 30Q chromatography of BB subunits: lane 1 is the Low molecular weight (LMW) marker; lane 2 is sucrose sample; lane 3 is periplasmic fraction; lane 4 is S30Q fraction 9; lane 5 is the S30Q fractions (18 to 21); lane 6 is the S30Q fraction (25); lane 7 is the S30Q fraction (27); lane 8 is S30Q fraction (32) and lane 9 is S30Q fraction (35); and lane 10 is the waste fraction. 34

Figure 6A. Source 30 Q chromatography results for BA subunits: absorbance at 280 nm, and peak is between fraction numbers 14 and 26. 34

Figure 6B. SDS–PAGE results from peak fraction of the Source 30Q chromatography of BA subunits: lane 1 is the Low molecular weight (LMW) marker; lane 2 is sucrose sample; lane 3 is periplasmic fraction; lane 4 is S30Q fraction (15); lane 5 is the S30Q fraction (18); lane 6 is the S30Q fraction (20); lane 7 is the S30Q fraction (32); lane 8 is S30Q fraction (35); lane 9 is the waste fraction; and lane10 is the flow throw fraction. 35

Figure 7A MonoQ chromatography results for BB subunits: absorbance at 280 nm and peak is between fraction numbers 31 and 38. 36

Figure 7B. SDS–PAGE results from peak fraction of the MonoQ chromatography of BB subunits: lane 1 is the Low MW marker; lane 2 is periplasmic fraction; lane 3 is MonoQ Fraction 33; lane 4 is MQ fraction (35); lane 5 is the MQ fraction (38); and lane 6 is the MQ fraction (42). 36

Figure 8A. MonoQ chromatography results for BA subunits. Absorbance at 280 nm . Peak between fractions numbers 21 and 28. 37

*Figure 8B.*SDS–PAGE results from peak fraction of the MonoQ chromatography of BA subunits: lane 1 is the molecular weight marker; lane 2 is periplasmic fraction; lane 3 is S30Q sample; lane 4 is MQ fraction (22); lane 5 is the MQ fraction (23); lane 6 is the MQ fraction (26); lane 7 is the MQ fraction (27); lane 8 is MQ fraction (32); and lane 9 is the flow through fraction. 37

*Figure 9A.*Gel Filtration chromatography results for BB subunits: absorbance at 280 nm and peak is located between fraction numbers 16 and 21. 38

Figure 9B. SDS–PAGE results from peak fraction of Gel Filtration chromatography of BB subunits: lane 1 is the MW marker; lane 2 is periplasmic fraction; lane 3 is the GF fraction (19); and lane 4 is the GF f21. 39

Figure 10A. Gel Filtration chromatography results for BA subunits: absorbance at 280 nm and peak between fraction numbers 19 and 24. 39

Figure 10B. SDS–PAGE results from peak fraction of Gel Filtration chromatography of BA subunits: lane1 is the MW marker; lane 2 is periplasmic fraction; lane 3 is S30Q sample; lane 4 is MQ sample; lane 5 is the GF fraction (19); lane 6 is the GF fraction (20); lane 7 is the GF fraction (21); lane 8 is GF fraction (22); and lane 9 is the GF fraction 24. 40

Figure 11. Circular Dichroism (CD) results of chemical denaturing of the four subunits. 41

Figure 12. CD results of four subunits**.** Melting denaturing of AB-AA subunits at 208 nm showed in part A at left side. Melting denaturation of BB-BA subunits at 215 nm showed in part B at right side. 42

Figure 13. Fractions folded and unfolded percentage of the four subunits; AA, AB, BB and BA as function of Temperature (K). Calculated free energy for un-folding at Melting Temperature is provided. 43

Abbreviations

CD Spectroscopy	Circular Dichroism Spectroscopy
CFAs	Colonization factor antigens
CSs	Coli surface antigens
DSC	Donor strand complementation
DSE	Donor Strand Exchange mechanism
EB	Elution buffer
ETEC	Enterotoxigenic Escherichia coli
Fn	Fibronectin
GdnHCl	Guanidine Hydrochloride
Ig	Immunoglobulin
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
LT	Heat-labile enterotoxin
PCFs	Putative colonization factors
PCR	Polymerase chain reaction
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
ST	Heat-stable enterotoxin

1 Introduction

1.1 Background

Enterotoxigenic *Escherichia coli* (ETEC) is one of the most important causative organisms of diarrhoea between people worldwide (Jiang, Lowe et al. 2002; Rao, Wierzba et al. 2005; Tobias, Holmgren et al. 2010). It is characterized by high morbidity and mortality among infants and children less than five years old in the developing countries (Tobias, Holmgren et al. 2010). About 1.5 million deaths occur per year due to acute infectious diarrhoea for children of developing world caused mainly by ETEC, rotavirus, *Vibrio cholera*, and *Shigella* spp (Qadri, Svennerholm et al. 2005). ETEC is regarded as the second most common cause of diarrhoea after retrovirus in children under 5 years old (Rivera, Ochoa et al. 2010), which causes about 400,000 deaths among infants and young children, and about 7.5 million cases of gastroenteritis worldwide per year (Guerena-Burgueno, Hall et al. 2002; Gupta, Keck et al. 2008). In addition to the life threatening and death of infants worldwide, the predisposing to ETEC infection as a single pathogen for children for one or more times in the childhood period, will cause malnutrition and growth stunted by 2 years of age when compared with those children of no ETEC infection (Qadri, Saha et al. 2007).

Adult visitors and tourists from North America and Europe to the endemic areas of the developing countries are often affected with the traveller's diarrhoea caused mainly by ETEC (25% - 45%) and carry back the pathogen to their home lands (Jiang, Lowe et al. 2002; Qadri, Svennerholm et al. 2005; Cohen 2010). Military persons who travel to and reside in these countries are facing the same problem (Cohen 2010). For example, 57% of the American troops settled in Middle East during 1990 to 1991 suffered from diarrhoea and about 20% of the units that settled in Saudi Arabia lost their duty due to sickness, meanwhile ETEC and *Shigella* were the major causative agents (Guerena-Burgueno, Hall et al.

2002). Diarrheal is the most challenging medical problem facing the military units that travel to such endemic regions, where ETEC is the most common causative agent (Cohen 2010).

1.2 Role of Environment and Social factors

ETEC remains endemic around the year but shows highest activity during warm seasons, which reflect the seasonality of the ETEC outbreaks and infections. This seasonality suggests that the tourists and travellers are more susceptible to illness during these times (Qadri, Svennerholm et al. 2005). Different phenotypes of ETEC are found according to the different geographical areas, for example the ST-only ETEC was more isolated from Kenya and represented 51% (Shaheen, Kamal et al. 2003) of the studied clinical cases, LT-only ETEC was most common in Jamaica with 58%, (Jiang, Lowe et al. 2002) and the, LT/ST ETEC was the predominant cases in India that represented about 45% (Jiang, Lowe et al. 2002) of the studied cases.

1.3 Risk factors of ETEC infections

The main source for infection with ETEC is the faecal-oral route from contaminated food, water and beverages due to poor hygienic conditions (Yates 2005). For example the contaminated hands of food handlers or contaminated drinking water and food resources. Regions of rapid population growth especially in the suburban regions served with poor quality water, poor infrastructure facilities and lack of sewage systems will increase the risk factors of ETEC infections (Gupta, Keck et al. 2008). Some environmental events can also enhance the risk and the epidemiology of ETEC; for example, during floods in Bangladesh the ETEC was a major cause of epidemic diarrhoea (Abu-Elyazeed, Wierzba et al.1999). In addition, serving of the water and weaning food to infants of about 3 months of age is reported to cause ETEC infection (Gupta, Keck et al. 2008).

1.4 Bacteriology

1.4.1 Classification and serotypes

Enterotoxigenic *Escherichia coli* (ETEC) belongs to the Enterobacteriaceae family, which is mostly Gram-negative, non-spore-forming, motile, lactose fermentative rod shaped organisms. ETEC can be classified further according to the serotype group like; O serotype (lipopolysaccharide antigen), H serotype (represent the flagellar antigen) and K serotype (capsular antigen) (Gupta, Keck et al. 2008).

1.4.2 Virulence factors and CFs – binding sites and their effect

In order to induce the disease, the ETEC strains should attach or colonize the small intestine mucosa after being ingested by individuals. This attachment to the host enteric epithelial cells is facilitated by the presence of surface antigens called colonization factor antigens (CFAs), coli surface antigens (CSs), or putative colonization factors (PCFs). These factors are mostly immunogenic proteins that bind to receptors on the host intestinal mucosa. Most of the CFs are fimbrial or fibrillar proteins, while some are not fimbrial in structure, for example CFA/I is a fimbrial antigen whereas CFA/II and CFA/IV consist of a wide variety of antigens (Tobias, Lebens et al. 2008; Rivera, Ochoa et al. 2010).

Up to date, more than 25 human colonization factors (CFs) are known for ETEC. The major CFs where frequently seen recently in clinical isolates were the CFA/I, CS3 and CS6 (Guerena-Burgueno, Hall et al. 2002). The CS6 antigen is taking up growing interest since it is isolated from clinical specimens from many countries and frequently seen worldwide (Byrd and Cassels 2003; Ghosal, Bhowmick et al. 2009; Rivera, Ochoa et al. 2010). Other main virulence factors for ETEC that causes secretory diarrhoea in the affected persons are the production of heat-labile enterotoxin (LT) and/or heat-stable (ST) enterotoxin that secreted after adhesion to the host intestinal mucosa by the help of CFs (Qadri, Saha et al. 2007).

LT enterotoxin is an immunogenic protein multimer encoded by the *eltA* and *eltB* genes, and consists of one A subunit surrounded by five identical B subunits. The A subunit is responsible for the catalytic activity (enzymatic) of the toxin, while the B subunits are responsible for the irreversible binding activity to the host mucosal cells surfaces gangliosides GM1 and GD1b. LT toxin resembles cholera toxin in many aspects like structure, physiology and mode of action or pathogenesis process. The A subunits stimulate the cyclic AMP (cAMP) and increase its production by binding to intracellular adenylyl cyclase. The other heat-stable (ST)

enterotoxin is non-immunogenic small peptide encoded by two genes: *estA* and *st1*, that produce STh and STp, respectively. The STh is isolated from human and the STp is isolated from pigs. STh binds to and activates the membrane guanylate cyclase resulting in increased intracellular cGMP while STp increases the cytosolic Ca^{2+} concentration and activates the release of prostaglandin E_2 and serotonin. The increase and activation of both cAMP and cGMP induced by LT and ST enterotoxin, respectively, lead to increased chloride secretion and ions from the crypt cells and impairment of the absorption of sodium chloride pump causing watery diarrhoea (Sears and Kaper 1996; Kaper, Nataro et al. 2004).

1.5 Clinical Feature

Most clinical cases of diarrhoea caused by ETEC can occur early after eight hours from the infection and last about four days without treatment (self-limited), while symptoms vary from mild to severe (Rivera, Ochoa et al. 2010). The onset of illness can be varied, shorter or longer (Gupta, Keck et al. 2008). The symptoms of the illness range from mild to severe. The common feature of this illness is characterized by the sudden onset of acute watery diarrhoea that can be accompanied by vomiting and abdominal cramps. Dehydration, dry mouth, rapid pulse and decreased blood pressure occur in severe cases when large volumes of body fluids and electrolytes are lost in the stool, which can develop to shock. Fever and bloody stool are usually not seen. When hydration is administrated, the patient will survive without any sequels (Qadri, Svennerholm et al. 2005; Gupta, Keck et al. 2008; Rivera, Ochoa et al. 2010).

The pathophysiology of the ETEC infection resembles the disease caused by *Vibrio cholera* (causes cholera disease) and the symptoms are usually identical especially in adult patients. The secretory diarrhoea is formed after secretion of the (LT) and/or (ST) in the intestinal lumen that leads to increased chloride secretion from the enterocytes and thus increasing osmotic pressure in the gut lumen. The result is accumulation of large volume of fluids that will be secreted out from the body into the stool causing progressive and large amounts of fluids and electrolyte loss.

1.6 Prevention and Vaccination

Food supplementation with zinc for infants and children less than 5 years old will help to protect them from the malnutrition and enhance their ability to resist the ETEC infection and other enteric pathogens such as cholera and *Shigella* by stimulating adaptive immune response (Albert, Qadri et al. 2003). Other important fac-

tor is the breast feeding that can reduce the ETEC diarrhoea and mortality in the infants(Giugliani and Victoria 2000).Maternal milk contains secretory immunoglobulin (Ig A) antibodies and anti-inflammatory factors. But the breast milk is regarded limited factor since when stopped and after weaning there will be a higher risk for infection since the transferred maternal immunity is stopped (Qadri, Svennerholm et al. 2005). The behavioural and social factors will exaggerate these risk factors of infection by supporting the weaned infants with contaminated food or water. For example in Mexico, people used to give a barley drink to infants about 3 months of age which increases the incidence of infectious diarrhoea caused by ETEC (Long, Wood et al. 1994). Personal hygiene is regarded more important protective factor against ETEC diarrhoea in Egypt (Abu-Elyazeed, Wierzba et al. 1999).

1.7 CS6 gene of ETEC

CS6 is one of the most frequently isolated antigens from clinical samples in various parts of the world and getting higher attention in research and study regarding the pathogenesis process of ETEC (Nicklasson, Sjoling et al. 2008). CS6 antigen belongs to unique antigen group that is non-fimbrial in structure. The genes belonging to the CS6 follow a typical bacterial operon structure, which consists of four genes *CssA, CssB, CssC* and *CssD,* which is expressed and transcribed as a single message in low amounts and transported to the cell surface chaperon-usher pathway (Tobias, Lebens et al. 2008; Sabui, Ghosal et al. 2010).

The *CssA* and *CssB* genes are encoding for two main heterologous structural subunits CssA and CssB, respectively. The *CssC* gene encodes for the chaperone protein, and finally the *CssD* gene encodes for the usher protein. The chaperon protein (CssC) is assumed to be responsible for folding of the structural subunits (CssA and CssB) and help in their protection from enzymatic proteolysis in periplasm. The usher protein (CssD) transports the structural subunits (CssA and CssB) from chaperon across the cell membrane to the cell surface (Ghosal, Bhowmick et al. 2009; Jansson, Tobias et al. 2009). The production of two subunits CssA and CssB is assumed to be 1:1 that means equal amounts are produced of these both subunits concurrently according to previous study (Ghosal, Bhowmick et al. 2009)

It is now established that the CS6-expressing ETEC binds to human small intestine epithelial cells through the binding of CssA subunit to the extracellular protein matrix, fibronectin (Fn) in CFU-dependent saturable manner. And the ability of CS6 to bind to glycosphingolipid as adhesion receptors is achieved through the CssB subunit which have a highly specific binding capacity to the Sulfatide

(SO3-3Galb1Cer), the main acid glycosphingolipid of the human small intestinal epithelial cells (Jansson, Tobias et al. 2009; Sabui, Ghosal et al. 2010). The immunologic characteristic of the structural subunits of CS6 gene to induce immunity is analysed and found that the purified CssA subunit have a poor immunogenic ability, while the CssB subunit is found to be highly immunogenic (Ghosal, Bhowmick et al. 2009).

1.8 Chaperone/usher pathway and Donor strand complementation theory

Chaperone/usher pathway is one major method for the assembly of the adhesive organelles of gram-negative bacteria, Enterobacteriaceae family. It is a simple mechanism that has only two specific proteins, a periplasmic chaperone and an usher protein used for assembly and transportation of the folded subunits (Zavialov, Batchikova et al. 2001; Zavialov, Kersley et al. 2002; Zavialov, Zav'yalova et al. 2007).

There are two families of the periplasmic chaperones classified according to sequence analysis, the FGS and FGL. FGS chaperones have short F1-G1 loop, which are used mainly for rigid pili assembly, while FGL chaperones have long F1-G1 loop and are mainly dedicated for assembly of non-pilus (fimbrial) organelles (Zavialov, Zav'yalova et al. 2007; Zavialov and Knight 2007). The periplasmic chaperones roles are to help protecting the formed periplasmic subunits by stable binding and avoiding the subunits from unwanted polymerization and proteolytic enzymes by covering their assembly surfaces. The final role of the periplasmic chaperones is to transport the subunits to the outer membrane molecular usher (Zavialov, Kersley et al. 2002; Remaut and Waksman 2004; Zavialov and Knight 2007). The outer membrane molecular usher acts by releasing the subunits from the chaperone and help in their assembly or polymerization. Finally, molecular usher transports the formed fibres to the cell surface (Zavala, Zav'yalova et al. 2007).

1.9 Chaperone structure and interaction with subunits

The periplasmic chaperones are formed of two immunoglobulin-like (Ig-like) domains forming an L-shaped or boomerang-shaped, molecule (Remaut and Waksman 2004; Busch and Waksman 2012; Waksman and Hultgren 2009). The pilin or fimbriae have a hydrophobic cleft or groove on their surface due to their structure building, where they lack the seventh C-terminal ß-strand and thus form incomplete Ig-like structure. The chaperone–subunit attachment is done by

donating the G1 ß-strand of the chaperone to the pilin (fimbriae) hydrophobic groove to form a stable complex. This process in which the chaperone and subunit are attached to form the stable complex is called donor-strand complementation **(**DSC) (Remaut and Waksman 2004; Busch and Waksman).

The pilin (fimbriae) subunits are attached to each other forming polymerizations occurs through Donor Strand Exchange mechanism (DSE). In this mechanism the chaperone G1 donor ß-strand is replaced by N-terminal extension of the next subunit in assembly to form the polymer (Busch and Waksman 2012).

The idea of this study is based on the donor strand complementation mechanism (Zavialov, Batchikova et al. 2001; Zavialov, Kersley et al. 2002; Zavialov, Zav'yalova et al. 2007) where the structural subunits (here CssA and CssB) are assumed to be attached to each other's similarly to their attachment within classical chaperone usher pathway through the usher G1 donor ß-strand, in which, the chaperone G1 donor ß-strand is replaced by N-terminal extension of the next subunit in assembly (CssA and CssB) by the DSE mechanism (Busch and Waksman 2012).

According to the DSC mechanism we will try to find the assembly pattern of the CssA and CssB based mainly on Caf1 project (Zavialov and Knight 2007). The schematic versions of the complemented subunits are shown in Figure 1 below.

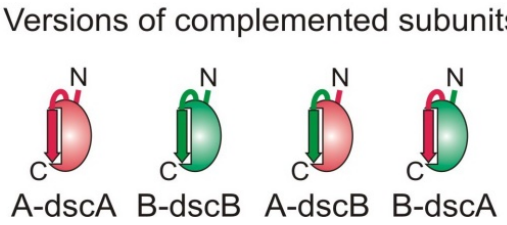


Figure 1. Versions of complemented subunits

We expect to have three possible polymers to be formed according to this study shown in Figure 2 below as following: Homo-polymerization of subunit A (M1), homo-polymerization of subunit B, (M2) and the hetero-polymerization of alternating both subunit A and subunit B together (M3)

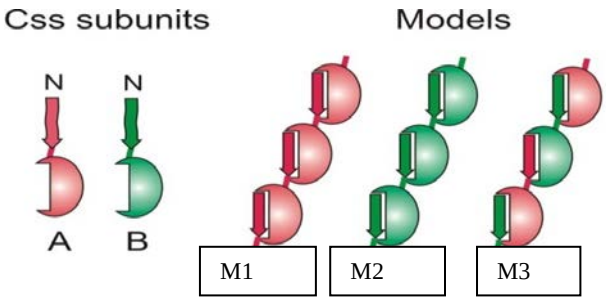


Figure 2. The possible polymers of this study.

The aim of this study is to put a step forward revealing the structural assembly of the CS6 expressed proteins, which help more in understanding the pathogenesis, diagnosis and future treatment methods. Vaccination as well can be studied and developed as a preventive method against ETEC.

Note: In this thesis

- CssA_dscA is simplified to A-dscA or AA,
- CssA_dscB is simplified to A-dscB or AB,
- CssB_dscB, is simplified to B-dscB or BB,
- CssB_dscA is simplified to B-dscA or BA.

2 Materials and methods

2.1 Expression Plasmids Construction

The genes of the four subunits (*CssA_dscA, CssA_dscB, CssB_dscB, and CssB_dscA*) were already prepared inside the vector (pUC57) by Invitrogen Co. The (pUC57) vector is surrounded by the EcoRI and SacI restriction sites, forming the PUC57-CssAA, PUC57-CssBB, PUC57-CssAB and PUC57-CssBA.

For the Expression plasmid creation, the pET101D vector (Invitrogen) was used to create the pET-CssA_dscA, pET-CssB_dscB, pET-CssA_dscB and pET-CssB_dscA. pET101D vector have the same restriction sites (the *EcoRI* and *SacI*) as in pUC57 vector. Required expression plasmids were made by ligation of the EcoRI- SacI fragments from pUC57 vectors that contain our required genes to the same location of EcoRI- SacI fragments in the pET101D vector. Our target genes were located downstream of the T7 promoter.

Note: all the 4 subunits (AA, AB, BB, and BA) were treated the same in all the following methods in this work.

2.1.1 Isolation of the (*AA, BB, AB* and *BB*) DNA fragments by PCR

The four subunits genes were isolated using the PCR method, where the enzyme used was pfu Ultra DNA polymerase. All the ingredients of the PCR reaction including reaction mix and PCR program used in this experiment are shown in Table 1 and Table 2, respectively in the next page.

Table 1. *PCR reaction mixture ingredients used for AA and BB fragment isolation*

Ingredients	Composition in µl
Sterile Water	19.5
10x Pfu buffer	2.5
dNTPs (10 mM)	0.5
Forward primer (20 ρmol)	0.5
Reverse primer (20 ρmol)	0.5
Template DNA (10 ng/µl)	1.0
Pfu Ultra enzyme	0.5
Total reaction volume	25 µl

Table 2. *PCR program*

Steps	Cycles	Temperature and Time
Initial Denaturation- window1	1	95°C for 1 min
Denaturation- window 2	25	95°C for 1 min
Annealing	25	52°C for 1 min
Elongation	25	72°C for 8 min
Final elongation –window 3	25	72°C for 10 min
Hold in- window 4	1	4°C

2.1.2 Agarose Gel Electrophoresis

1% Agarose gel was prepared for analysing the resulting PCR product as follows: 1 gram of pure Agarose (Invitrogen) powder was weighed in 500 ml bottle, then added 100 ml 1x TAE buffer and heated in microwave till melting of the Agarose, then cooled down to 50 °C. After that, 50 ml of the cooled Agarose gel was taken and added 4 µl GelRed (Biotium-USA) then mixed properly. The gel was then poured in the gel-casting tray. Fresh 1x TAE buffer then added to the tank to be used as running buffer. Sample volume was 60 µl that loaded to the gel by using 50 µl PCR products and 10 µl 6x loading dye. DNA marker, 1000 bp, was loaded at the first well then the PCR products samples were loaded. Electrophoresis was run at 100 volts for 40-45 minutes. Finally, the gel was examined under the UV light shortly and photos were taken.

2.1.3 DNA Purification from Agarose gel

PCR products or DNA bands were cut carefully using clean scalpel under the UV lamp and gel slices were made small as possible. The purification was done using QlAquick Gel Extraction Kit Protocol (QIAGEN, USA) where the gel piece was weighed then added 3 volumes of buffer QG to one volume of the gel. Then the

mixture was incubated at 50 °C for 10 minutes with mixing each 2-3 minutes to dissolve the gel completely. Dissolved gel was then applied to the QIAquick spin column and centrifuged at 13,000 rpm for 1 minute to bind DNA, then the flow-throw was discarded, after that washed by (0.75 ml) Buffer PE and centrifuged again for 1 minute at 13,000 rpm. Flow-throw was discarded and centrifuge applied again with the same previous conditions. DNA elution was done by adding 50 µl Elution buffer (EB) to the centre of the column and centrifuged finally as before.

2.1.4 Cloning and ligation into pET101/ TOPO vector

This step was done for the construction of pET101-*CssA_dscA*, pET101-*CssB_dscB*, pET101-*CssA_dscB* and pET101-*CssB_dscA* using the TOPO cloning reaction. TOPO Cloning Ligation reaction mixture was prepared by taking 1 µl of fresh purified DNA from the previous step, salt solution 1 µl, sterile water 3 µl, and 1 µl of TOPO vector (Invitrogen, USA). The reaction was mixed gently in eppendorf tube and incubated for 20 minutes at room temperature.

Transformation of DH5α competent cells was done as follows: One vial of DH5α competent cells (Invitrogen, USA) were used per transformation reaction by thawing in ice for 5 minutes. Then 2 µl of the ligation mixture was added to the DH5α competent cells and incubated for 30 minutes inside ice. Then heat shock was applied by incubating the mixture directly in 42 °C water bath for 1 minute without shaking. Then removed back quickly and incubated again in ice for 5 minutes. 400 µl of fresh LB medium then added to the vial and incubated at 37 °C for 1 hour and shaking of 200 rpm. Finally, 400 µl of the transformation reaction is plated on Pre-warmed LB agar plates containing certain antibiotic (ampicillin 100µg/ ml) and spread well, and then the plates were inverted after absorption the solution and incubated at 37 °C overnight.

2.1.5 Colony PCR

This method was used for screening of positive transformation of the previous step. Some bacterial colonies were taken from the overnight culture from the LB agar plate and added to 4 ml LB broth containing appropriate antibiotic to our cell type (ampicillin 50µg/ ml), and incubated at 37 °C for about 3 hours under shaking at 200 rpm. Then 10 µl of the culture was incubated at 95 °C for 10 minutes using PCR heating block, then sample was immediately put in ice. This sample was our PCR template for the following Colony PCR. The Taq DNA polymerase primer (New England Bio lab, USA) and T7 forward primers used in this work. Table 3

and Table 4 show the Colony PCR mixture and Colony PCR program, respectively.

Table 3. *PCR reaction mixture of Colony PCR*

Ingredients	Composition in µl
Sterile Water	19.8
10x Taq buffer	2.5
dNTPs (10 mM)	0.5
Forward primer (20 ρmol)	0.5
Reverse primer (20 ρmol)	0.5
Template DNA (15 to 20 ηg/µl)	1.0
Taq DNA polymerase	0.2
Total reaction volume	25

Table 4. *Colony PCR program*

Steps	Cycles	Temperature and time
Initial Denaturation- window1	1	94°C for 2 min
Denaturation- window 2	25	94°C for 1 min
Annealing	25	50°C for 1 min
Elongation	25	72°C for 1 min
Final elongation –window 3	1	72°C for 2 min
Hold in- window 4	1	4°C

2.2 Plasmid Purification

The plasmid from the transformed cells that carries our insertion was purified using the QIAprep Spin MiniPrep Kit (QIGEN, USA). Firstly, about 600 µl of the overnight culture were put in a falcon tube of 15 ml size, centrifuged for 10 minutes at 4500rpm.Then the supernatant was discarded and cell pellets were suspended in 250 µl of the re-suspending buffer P1 and transferred to a microcentrifuge tube. After that 250 µl of P2 buffer was added and mixed thoroughly for 4-6 times. Then 350 µl of neutralisation buffer N was added and mixed the same as before, then centrifuged for 10 minutes at 13,000 rpm in microcentrifuge.

Supernatant was then transferred into QIAprep spin column and centrifuged for 1 minute at 13,000 rpm again. Flow-through was discarded, followed by washing the column by adding 0.75 ml of the washing buffer (PE) and centrifuged for 1 minute at 13,000 rpm. Flowthrough was discarded again and other centrifugation step was done as before to remove the residues of washing buffer. The QIAprep column is put in new clean microcentrifuge and the lid was cut, then the DNA was eluted by adding 50 µl of sterile water in the centre of the spin column, incubated

for 2 minutes then centrifuged for 1 minute. The solution was then tested by Nano drop and transferred to new eppendorf tube and the plasmids are stored at -20 ℃.

2.3 Glycerol stock preparation

After finishing from colony PCR and be sure that the transformed cells gave positive colonies, a stock solution from these positively clones of transformed cells was prepared by adding 200 µl of sterile filtered 88 % glycerol to 600 µl of the overnight culture, both in eppendorf tube. The solution was then shaken by vortex for 10 seconds and instant freeze was applied using liquid Nitrogen. The final glycerol stock is stored at -80 ℃.

2.4 Transformation of plasmids into expression cells BL21-DE3

Purified plasmid from previous plasmid purification step was used to be transformed into a homemade cell line (BL21-DE3) by heat shock method. Firstly the BL21-DE3 cells were taken from -80 ℃ and thawed in ice, and then 1 ul from purified plasmid was added, tapped and incubated for about 30 minutes on ice. After this, incubation in 42 ℃ water bath for 1 minute (as homemade cells) and transferred to ice for 5 minutes. Then 400 µl of LB medium was added and incubated at 37 ℃ for 1 hour while shaking by 200 rpm. The cell culture then spread on the inverted pre-warmed agar plate overnight at 37 ℃.

2.5 Screen of clones by test expression

This test was performed to test the expression of the positive clone of the BL21-DE3 cells from previous transformation step. One positive single colony was taken and added to 5 ml LB medium containing ampicillin of 100 mg/ml. The culture was then grown overnight at 37 ℃ with shaking of 200 rpm. Next day inoculation was done by adding 4 ml of the overnight culture to a sterile 50 ml falcon tube containing 20 ml fresh LB medium with ampicillin (100 mg/ml). Culture was grown for 1 hour with shaking by 200rpm at 37 ℃.

Optical density was measured of the culture cells at wavelength of 600 nm to determine their density every hour, where the cells induction was carried out at density between 0.5 and 0.8 of cells. The main culture sample was divided into 2 samples, 10 ml each. Cells induction was done by adding 20 ul IPTG (isopropyl ß-D-1-thiogalactopyranoside) to get the final concentration of 1 mM to one sample

of 10 ml and incubated for 3-4 hours at 37 °C while the other 10 ml sample remained not induced as negative control.

2.6 Protein harvesting from periplasmic space

Cell harvesting was done by taking 4 ml of both cultures separately, and then centrifuged for 15 minutes at 4500 rpm, where cells present in the pellets. Supernatant was discarded and cell pellets were kept. Cell pellets were then re-suspended (dissolved) with 600 ul buffer containing (20 % w/v sucrose, 50 mM Tris-HCl pH. 8.0, and 5 mM EDTA) and then transferred in microcentrifuge. Ratio of buffer to cells was 4 ml buffer/1 gram pellet. Then mixed and transferred to new eppendorf tube with proper labelling and incubated on ice for 10 minutes. Then centrifugation step was done at 5000 rpm for 15 minutes. Supernatant was collected and labelled as **sucrose fraction** and kept at -20 °C in cold room to be used for SDS PAGE analysis.

Cell pellets after centrifugation were suspended again with 200 ul of buffer (4 ml/g) containing 5 mM $MgSO_4$ and protease inhibitor, then kept incubated on ice for further 10 minutes and immediately centrifuged for 5 minutes at 10000 rpm. The clear supernatant was collected, representing the **periplasmic fraction** that was put in eppendorf tube with proper labelling and kept in -20 °C in cold room for further SDS PAGE and other analysis.

2.7 SDS-PAGE (15%) for protein expression analysis

Both stacking and separation SDS –PAGE gels were prepared according to Table 5 below.

Table 5. *SDS PAGE ingredients for separation and stacking gels*

15% separating gel, pH 8.8	ml	4% stacking gel, pH 6.8	ml
88% Glycerol	0.525	H_2O	1.84
1M Tris-HCl, pH 8.8	1.87	1M Tris-HCl, pH 6.8	0.312
30% AA-bis AA (30:1)	2.5	30% AA-bis AA (30:1)	0.325
10% SDS	0.05	10% SDS	0.025
10% PSA	0.05	10% PSA (Acrylamide+Bis-acrylamide (29:1))	0.025
TEMED	0.005	TEMED	0.002
Total	5		2.5

Protein samples were prepared firstly by taking 30 ul of each sample in microcentrifuge tube then adding 10 ul of the 4x loading dye and 1 ul of β-mercaptoethanol (0.1 M) all in the same tube then boiled at 95 °C for 5 minutes. Meanwhile sample preparation, the gel was prepared by pouring separating gel and wait till solidified then some distilled water was added to make flat surface for about 1 cm below the well comb. Stacking gel was then poured and the comb was inserted directly to make the wells. When the stacking gel was polymerized, the comb was removed, samples were added to the wells and the running buffer was added. About 10 ul of the samples were added to wells and the low molecular weight ladder was loaded in the first lane. Electrophoresis was run at 200 V for 1 hour using 1x SDS Running buffer (25 mM Tris-base +192 mM Glycine + 0.1% SDS + H_2O) or till the tracking dye reached to end.

2.7.1 Gel staining with Coomassie Blue

The gel was removed carefully from the casting glass and put in clean plastic box then added the fixation solution (30 % ethanol and 10 % acetic acid) over the gel, then heated near boiling using the microwave for about 10-15 seconds, then the fixation solution was removed and stored to be used again. The staining Coomassie Blue solution was added to the gel, heated near boiling using microwave as before. Gel was then incubated within the Coomassie Blue solution for about 5 minutes under shaking at room temperature. After that the gel was washed with distilled water, heated for about 2 minutes near boiling by microwave, and incubated at room temperature for about 5 minutes or more under gentle shaking. Water was changed and the rinsing of the gel with water was repeated till the gel background was low and the protein bands could be easily seen. Finally, the gel was scanned to reserve the results.

2.8 Large culture and purification of protein

This large culture technique was used to make overexpression of protein for the further purification steps. An overnight 4 litres large culture was prepared (4 flasks with 1litre each) by adding for each 1 litre 960 ml LB medium + 40 ml from the overnight *AA,AB,BB, and BA* cultures. 960 ul of ampicillin was added with concentration of (100mg/ml). Culture was then incubated for 2 hours and 15 minutes at 37 °C with shaking (large shaker incubator) at 145 rpm to reach an optical density (O.D) between 0.6 and 0.8. Measuring O.D was done followed by induction using (1 ml/L) of IPTG (1M).

Cultures were incubated again for protein expression for 3 hours at 37 °C with shaking at 145 rpm, and then centrifuged at 4200 rpm for 30 minutes using 1 litre plastic bottles. Cells were then suspended using 20 ml buffer containing 20% w/v sucrose, 50 Mm Tris-HCl pH 8.0 and 5 mM EDTA then kept on ice for 10 minutes. Centrifugation was done for 20 minutes at 6800 rpm and the supernatant was saved in a separate container as the **sucrose fraction** to be analysed by SDS-PAGE later. While cells pellets were re-suspended using 15ml of (5 mM) MgSO4 and protease inhibitor (for osmotic shock) then kept on ice for 10 minutes, after that centrifuged for 20 minutes at 7500 rpm. The resulting supernatant is the **periplasmic proteins** that were collected for next dialysis step.

2.9 Dialysis and filtration of protein

Dialysis was done for protein samples in order to decrease the salt concentration (desalting) and for further buffer exchange techniques. The sample was put in a semipermeable dialysis membrane (SPECTRA/POR) and kept overnight inside the cold room. The dialysis buffer was selected according to the next used chromatographic techniques, where the same loading buffer of the chromatography used as the dialysis buffer with the same pH. Table 6 below shows a list of dialysis buffers used for each subunit purification step.

Table 6. *List of Dialysis buffer used*

Chromatography technique	Dialysis Buffer used	Elution Buffer used	Used for purification of
Anion Exchange-Source 30Q	20 mM Tris-HCl pH 7.4	20mM Tris HCl + 1M NaCl pH 7.4	AA,AB,BB and BA
Anion Exchange-Mono Q	20mM Bis-Tris pH 6.0	20mM Bis-Tris + 1M NaCl pH 6.0	BB and BA
Cation Exchange-Mono S	50 mM Sodium acetate pH 4.1	50 mM NaAc+ 1M NaCl pH 4.1	AA and AB
Gel Filtration	20 mM HEPES, pH 7.3	20mM HEPES +150mM NaCl pH 7.3	AA,AB,BB and BA

2.10 Protein purification by liquid chromatography

The following chromatography techniques were used for protein sample purification by using different columns by the automated system of ÄKTA Explorer, GE Healthcare, Sweden.

2.10.1 Anion-exchange chromatography (Source 30Q and MonoQ)

Source 30Q chromatography is used as the first liquid chromatography purification step for all the four subunits; AA, AB, BB, and BA. In this technique, the Source 30Q column was used as the first step of protein purification by anion-exchange chromatography. The column was washed with three column volumes (CV) of the elution buffer (20 mM Tris-HCl, 1M NaCl , pH 7.4) then equilibrated with three CV of loading buffer (20 mM Tris-HCl, pH 7.4). The dialysed and filtered protein samples were loaded at flow rate of 1 ml/ minute into the Source 30Q column. The column was then washed by using 1.5 CV of loading buffer to get the protein samples inside it. Elution step was done by using Elution buffer at a flow rate of 1 ml/ minute and gradient of 30% for 100 minutes. The SDS-PAGE analysis was used after this step in order to identify the fractions of purified protein and mainly the peak and its related fractions. The protein sample was then put in dialysis membrane again using proper dialysis buffer, the same as loading buffer for the next MonoQ chromatography technique. In this step the dialysis buffer was 20mM Bis-Tris pH 6.0.

2.10.2 Mono Q chromatography (8 CV, 5MPa)

Mono Q anion exchange chromatography was used as the second purification step for the subunits *BB* and *BA* only. Firstly, pump was washed at flow rate of 3 ml per minute, then column elution done with 5 CV where the elution buffer used was 20 mM Bis-Tris, 1M NaCl, pH 6.0, then equilibrated with loading buffer of 20 mM Bis-Tris, pH 6.0, as illustrated in Table 6 previously. Sample was then loaded and eluted using 4 CV at the same flow rate of 3 ml/minute with 1.5 ml fraction size. Then column was washed and equilibrated again. The resulted fractions of interest were run in SDS-PAGE. Protein samples were then concentrated using Vivaspin-6 column 5000 MWCO PES, GE Healthcare for the next uses in Gel Filtration step.

2.10.3 Cation exchange chromatography (Mono S)

Mono S column (1 ml) was used as Cation exchange chromatography second purification step for the subunits AA and AB only. The elution buffer used for washing was (50 mM Sodium acetate pH 4.1 +1 M sodium chloride) by 30 CV, then equilibrated by 30 CV of the loading buffer (50 mM Sodium acetate pH 4.1) with flow rate of 1 ml/min. Protein sample was loaded to the column in the same flow rate of 1 ml/min using 20 CV loading buffer then eluted at flow rate of 0.5 ml/min with fractions of 0.5 ml/tube. Gradient was 30 % for 30 minute for elution. The

SDS-PAGE analysis was used after this step and the protein samples then concentrated using Vivaspin-6 column 5000 MWCO PES (GE Healthcare).

2.10.4 Size exclusion chromatography (Gel filtration)

Gel filtration method was used to separate proteins of different sizes and can determine their molecular weight. In our experiment here we used the Superdex 75E 120 ml column (GE Health Care) to purify our proteins. The column was equilibrated with 130 ml of 20 mM HEPES, pH 7.3, 150 mM NaCl. Flow rate was 1 ml/minute. About 5.5 ml from Mono Q sample was concentrated to get approximately 1 ml (10.5 mg/ml) of concentrated sample. Elution rate was about 1 ml/minute and the fraction size was 2 ml. After pure protein elution is finished the SDS-PAGE was run.

2.11 Protein Stability study by Circular Dichroism (CD) Spectroscopy

Circular Dichroism is used to study many aspects of protein structural properties, like the proteins behaviour and operations under certain conditions and other structural changes either by chemical denaturation or temperature, which are all important to understand their functions. The main uses of CD are to study the secondary structure of proteins, tertiary structure fingerprinting, and integrity of cofactor binding sites. In CD Spectroscopy the difference between right and left circularly polarized light is measured as a function of wavelength. The far-UV bands (180-240 nm), which came mainly from peptide bond absorption, are reflecting the secondary structure of the protein like α-helix and β-sheet. While the near-UV bands (240-340 nm) of proteins are reflecting the tertiary structure of the protein, which derived mainly from peptide bonds absorption.

2.11.1 Chemical denaturation using Guanidine Hydrochloride (GdnHCl)

Different concentrations of GdnHCl were used to test the stability and unfolding of the four subunits (AA, AB, BB and BA) as shown in the Table 7 below. Protein subunits samples were diluted 4 times with 10 mM phosphate buffer, pH 7.0 and used different concentration of GdnHCl in buffer B solution. Samples were incubated overnight (16-21 hours) at 21°C, then test was run.

Table 7. *Reaction mix of GdnHCl and protein subunits by Circular Dichroism (CD)*

Sample	S1	S2	S3	S4	S5	S6	S7	S8	S9
GdnHCl in sample (M)	0	0.5	1	1.5	2	2.5	3	3.5	4
Protein in 10 mM PB, pH 7.0 (µl)	100	100	100	100	100	100	100	100	100
8M GdnHCl (µl)	0	12.5	25	37.5	50	62.5	75.0	87.5	100
H2O (µl)	100	87.5	75	62.5	50	37.5	25	12.5	0
Total (µl)	200	200	200	200	200	200	200	200	200

2.11.2 Thermal stability of the studied protein subunits

Diluted protein samples were used in the absence of GdnHCl, to determine temperature denaturation of the four subunits. Samples were heated from 30°C to 90°C at a heating rate of 1°C per minute, and the resulted wavelength was recorded continuously.

2.11.3 Free energy change ΔG (kJ/mol) calculation

The free energy change ΔG (kJ/mol) calculation was done here depending on the thermal denaturing curves of the CD results. The ellipticity over a range of temperature was normalized between 0-1 that gave the curve of a fraction (%) of protein in folded and unfolded state. This was done by using the following formula in Excel sheet:

(Y) - Min (distribution)/ (Max (distribution)) - (Min (distribution)).

Here, Y represents the folded and unfolded percentage value in the curve between 0 and 1 depending upon the equation (f_f) + (f_u) = 1. Where $\boldsymbol{f_f}$ represents the total folded proteins and $\boldsymbol{f_u}$ represent the total unfolded proteins. The results of these curves are shown in Figure 13 at page 43 (Creighton, 1997; Lee et al., 2005).

For calculating the Equilibrium Constant, Keq, the following equation is used

$K = f_u / f_f$ (1)

For fraction unfolded represented by the formula $f_u = (Y_f - Y)/(Y_f - Y_u)$ (2)

And the fraction folded represented by the formula $f_f = Y - Y_u/(Y_f - Y_u)$. (3)

Y is the observed value of protein folding at transitional state, f_u is the unfolded protein value, and f_f is the folded protein value. The *Keq* was calculated based on the points and data from the protein folding and unfolding percentage shown in Figure 13 in the Results section. For calculation used here, Y value was fixed for 0.5 as representing the transitional state where about 50 % of the proteins are in

folded state and the others are unfolded. The f_u and f_f values were calculated by choosing two points at least at folded state for f_f , and two points at unfolded state for f_u.

The free energy change, ΔG, is calculated using the following equation

$$\Delta G = - RT \ln K \quad (4)$$

Where R is the gas constant 8.314 (J /mol. K) and T is the absolute temperature in Kelvin and K is the calculated equilibrium constant (Creighton, 1997; Lee et al., 2005).

3 Results and Discussion

3.1 Construction of Expression Plasmid for the four complemented subunits

Required expression plasmids were constructed by ligation of the EcoRI-SacI fragments from pUC57 vectors, which contains our required genes, to the same location of EcoRI-SacI fragments in the pET101/TOPO vector. Figure 3 below shows the successful PCR results that confirmed the fragment isolation.

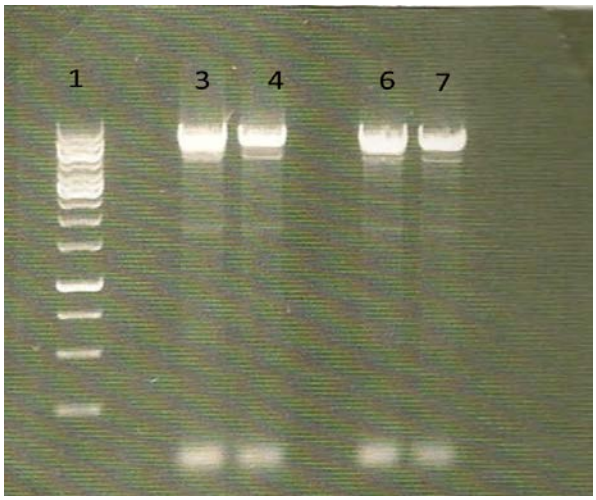


Figure 3. PCR results of 4 subunits isolation at 1 % Agarose gel: lane 1 represents the ladder (1000bp); lane 3, AA; lane 4, AB; lane 6, BB; and lane 7, BA.

The four subunits genes were then isolated and ligated to pET101/TOPO vector and then transformed successfully to DH5α competent cells. After that, Colony PCR was conducted to confirm the positive transformation of the competent cells. Results are not shown here.

3.2 Test expression results on SDS-PAGE

Test expression was done for the four subunits in BL21-DE3 cells to enhance the protein expression conditions and to select the clones of higher expression potential. The SDS-PAGE analysis (Figure 4) shows that all subunits were successfully expressed.

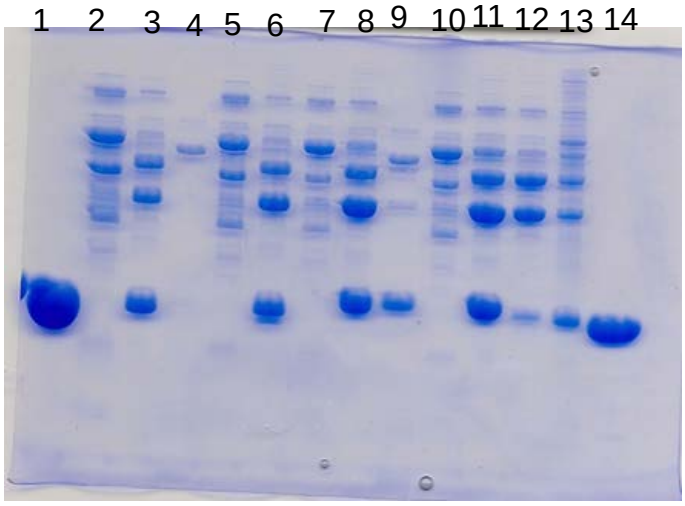


Figure 4. Test expression results on SDS-PAGE: lane 1, AB cells control; lane 2, AB cell (none induced); lane 3, AB cell (induced); lane 4, AB cell sucrose fraction; lane 5, AB old cell (none induced); lane 6, AB cells (induced); lane 7, BA cells (none induced); lane 8, BA cells (induced); lane 9, BA cells sucrose fraction; lane 10, BA old cells (none induced); lane 11, BA old cells (induced); lane 12, AA cells (induced); lane 13, BB cells (induced) and lane 14, control AB cells.

Old samples, newly induced, and non-induced samples were run. The results shown in Figure 4 indicate that the induced samples were expressed successfully for all the 4 subunits. Lane 3 shows a successful expression of the AB subunits and lane 8 shows the expression of BA subunits. The AA subunits expression was shown in lane 12 and successful expression for the BB subunits, was shown in lane 13. Here we noticed that the AB and BA expression was stronger than BB and AA expression.

Large cultures were done to get the overexpression of the protein successfully after the expression tests results. Then all the 4 subunits periplasmic proteins (AA, AB, BB and BA) were dialysed using the loading buffer of Source 30Q as an anion-exchange chromatography using the 20mM Tris-HCl pH 7.4

For the following protein purification steps, results to be shown will be for the two subunits BB and BA regarding protein purification. Other subunits AA and AB were purified in the same techniques except that MonoS was used as Cation-

exchange chromatography rather than MonoQ that was used for BB and BA subunits.

3.3 Anion-exchange chromatography (Source 30 Q) results

Results of Source 30Q (S30Q) chromatography are shown in Figure 5 for BB subunits, and in Figure 6 for BA subunits.

Figure 5 shows the results for BB subunits. Figure 5A below shows the chromatogram of the Source 30Q purification at 280 nm absorbance and Figure 5B in the following page shows the SDS gel run results of the fractions from the S30Q purification step. The Main Peak was located between fraction numbers 16 and 22. Lane 5 in Figure 5B represents the main peak as our target protein fraction when compared to the Sucrose and Periplasmic fraction bands.

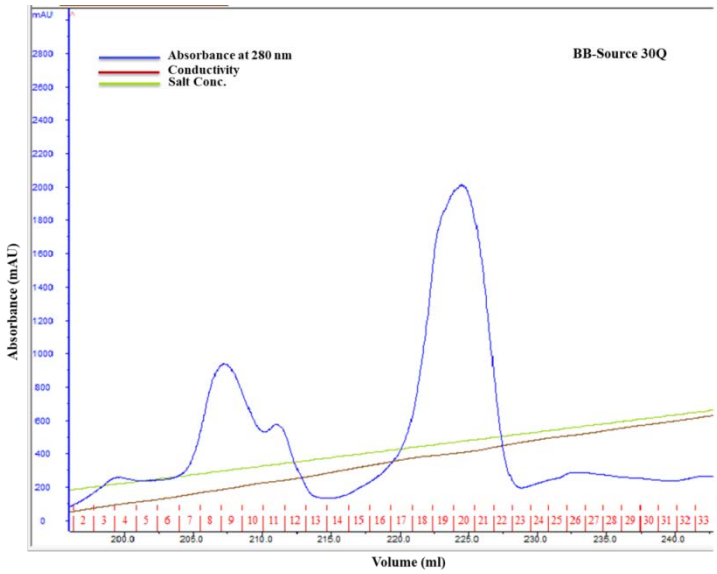


*Figure 5*A. Source30Q chromatography results for BB subunits: absorbance at 280 nm, and peak is between fraction numbers 16 and 22.

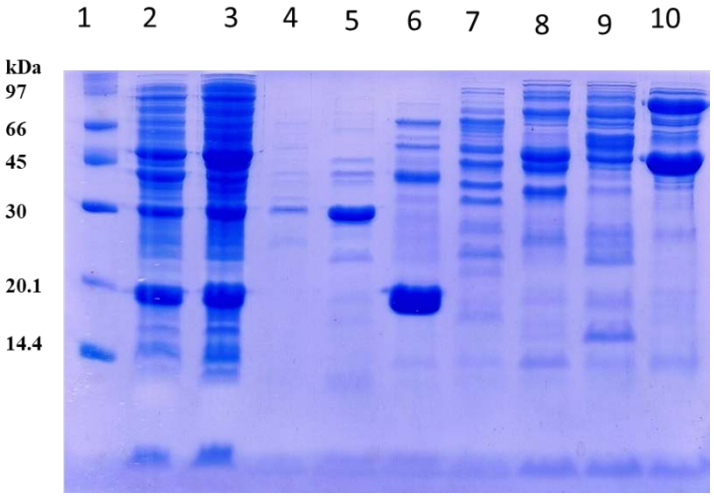


Figure 5B. SDS–PAGE results from peak fraction of the Source 30Q chromatography of BB subunits: lane 1 is the Low molecular weight (LMW) marker; lane 2 is sucrose sample; lane 3 is periplasmic fraction; lane 4 is S30Q fraction 9; lane 5 is the S30Q fractions (18 to 21); lane 6 is the S30Q fraction (25); lane 7 is the S30Q fraction (27); lane 8 is S30Q fraction (32) and lane 9 is S30Q fraction (35); and lane 10 is the waste fraction.

Figure 6 below shows the results for BA subunits. Figure 6A shows the chromatogram of the Source 30Q purification at 280 nm absorbance, and Figure 6B at the next page shows the SDS gel results of the fractions from the S30Q purification step. The peak was located between fractions number 14 and 26.

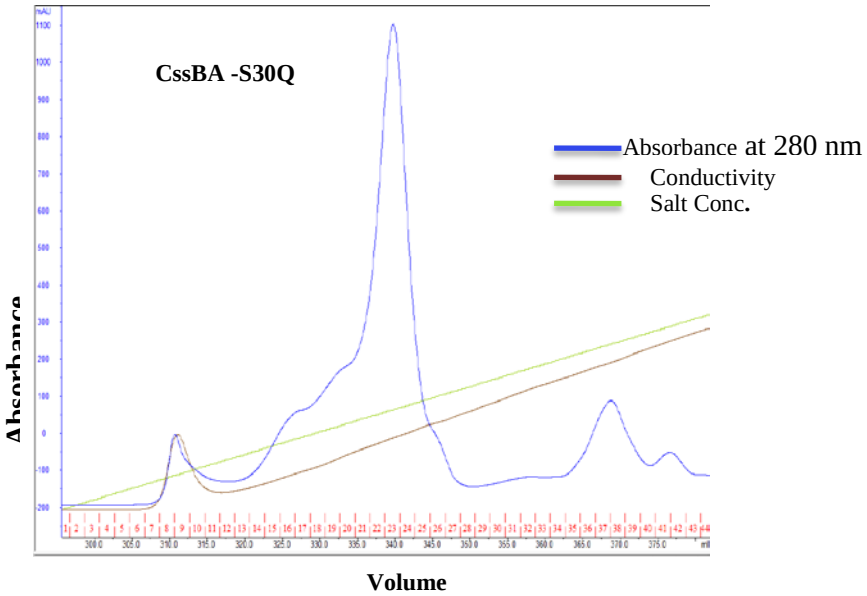


Figure 6A. Source 30 Q chromatography results for BA subunits: absorbance at 280 nm, and peak is between fraction numbers 14 and 26.

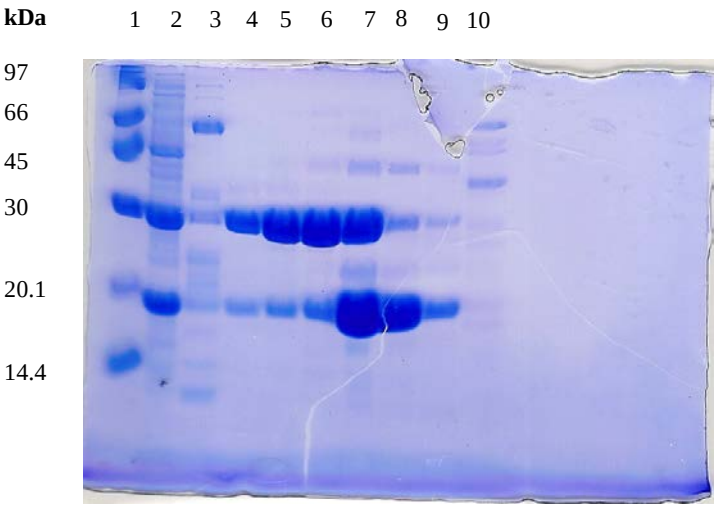


Figure 6B. SDS–PAGE results from peak fraction of the Source 30Q chromatography of BA subunits: lane 1 is the Low molecular weight (LMW) marker; lane 2 is sucrose sample; lane 3 is periplasmic fraction; lane 4 is S30Q fraction (15); lane 5 is the S30Q fraction (18); lane 6 is the S30Q fraction (20); lane 7 is the S30Q fraction (32); lane 8 is S30Q fraction (35); lane 9 is the waste fraction; and lane10 is the flow throw fraction.

3.4 Anion-exchange chromatography (Mono Q) results

Results of Mono Q chromatography are shown in Figure 7 and Figure 8 for BB and BA subunits. Figure 7A shows the chromatogram of the Mono Q purification at 280 nm absorbance, and Figure 7B shows the SDS gel results of the fractions from the Mono Q purification step. Peak was located between fractions number 31 and 38.

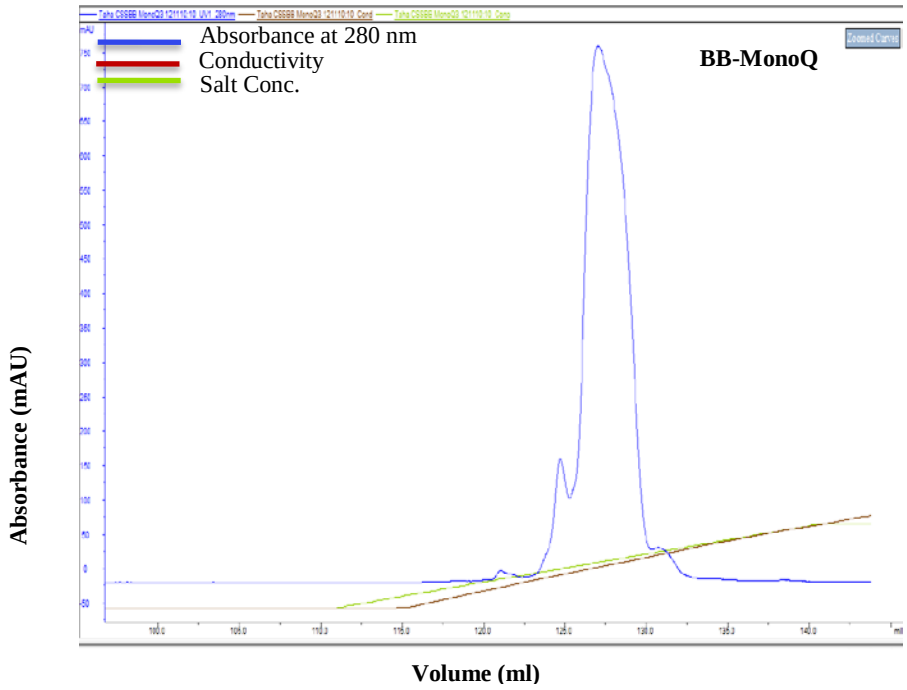


Figure 7A MonoQ chromatography results for BB subunits: absorbance at 280 nm and peak is between fraction numbers 31 and 38.

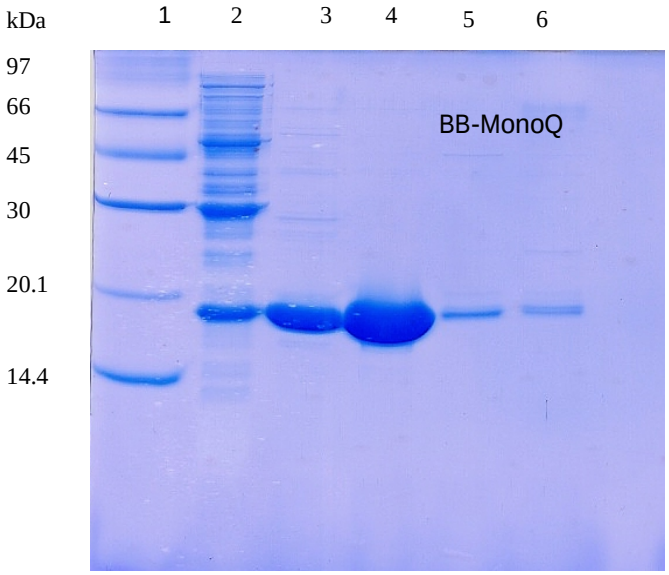


Figure 7B. SDS–PAGE results from peak fraction of the MonoQ chromatography of BB subunits: lane 1 is the Low MW marker; lane 2 is periplasmic fraction; lane 3 is MonoQ Fraction 33; lane 4 is MQ fraction (35); lane 5 is the MQ fraction (38); and lane 6 is the MQ fraction (42).

Figure 8 below shows the results for BA subunits. Figure 8A shows the chromatogram of the MonoQ purification at 280 nm absorbance, and Figure 8B shows the SDS gel results of fractions from the MonoQ purification step and previous S30Q peak results. Mono Q Peak was located between fraction numbers 21 and 28.

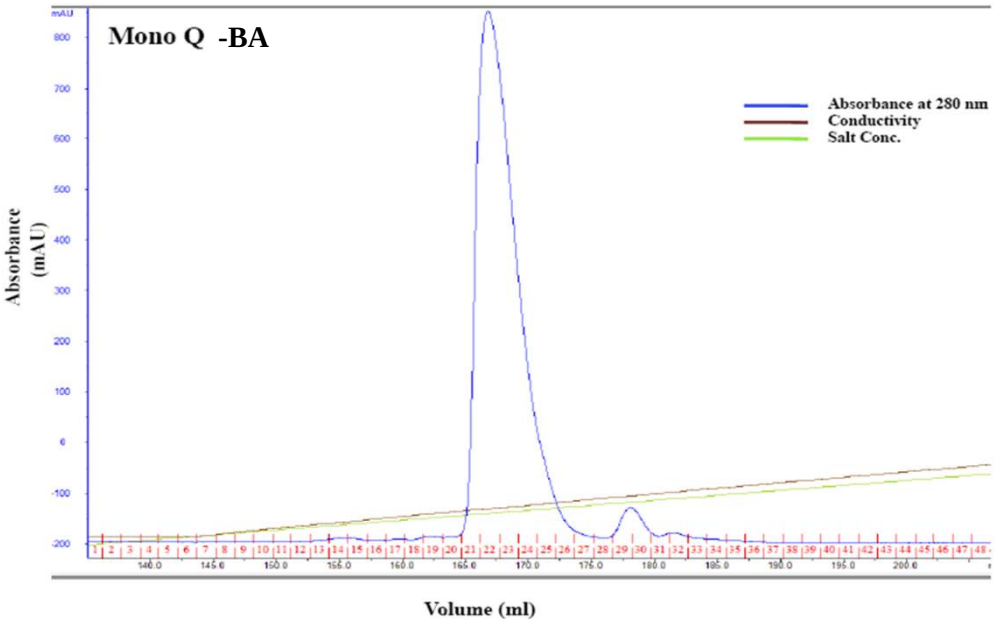


Figure 8A. MonoQ chromatography results for BA subunits: absorbance at 280 nm and peak is located between fraction numbers 21 and 28.

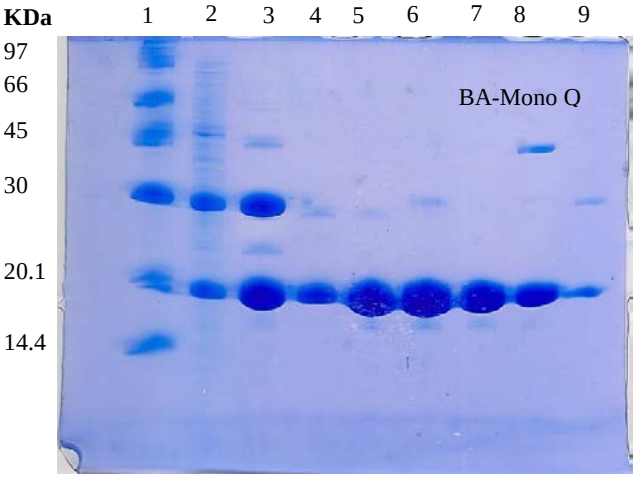


*Figure 8B.*SDS–PAGE results from peak fraction of the MonoQ chromatography of BA subunits: lane 1 is the molecular weight marker; lane 2 is periplasmic fraction; lane 3 is S30Q sample; lane 4 is MQ fraction (22); lane 5 is the MQ fraction (23); lane 6 is the MQ fraction (26); lane 7 is the MQ fraction (27); lane 8 is MQ fraction (32); and lane 9 is the flow through fraction.

3.5 Size exclusion chromatography (Gel filtration)

Results of Gel Filtration are shown in Figure 9 and Figure 10 for BB and BA subunits, respectively.

Figure 9 shows the results for BB subunits. Figure 9A shows the chromatogram of the gel filtration (GF) purification at 280 nm absorbance, and Figure 9B, at the next page, shows the SDS gel results from gel filtration purification step. GF Peak was located between fraction numbers 16 and 21.

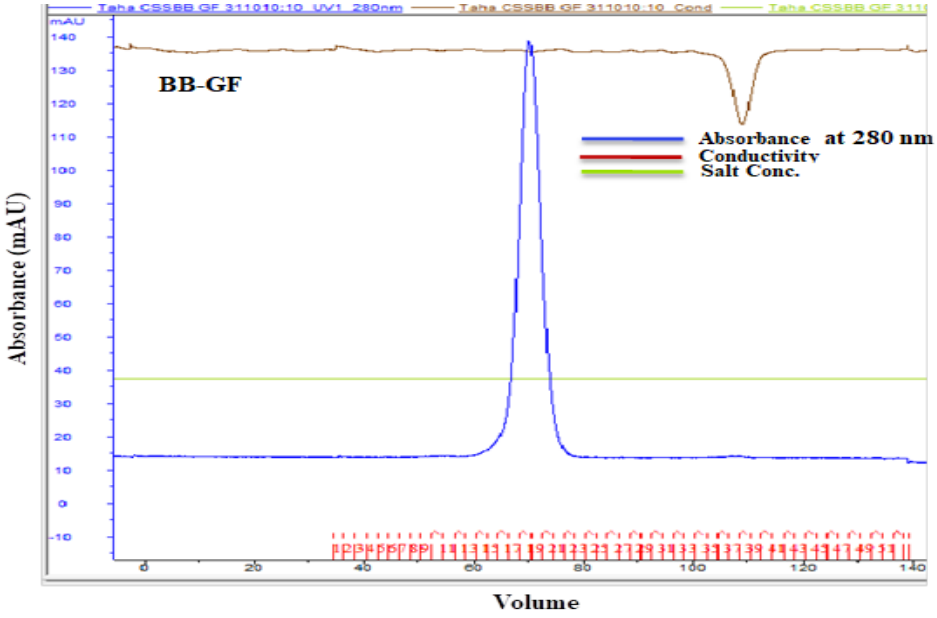


*Figure 9A.*Gel Filtration chromatography results for BB subunits: absorbance at 280 nm and peak is located between fraction numbers 16 and 21.

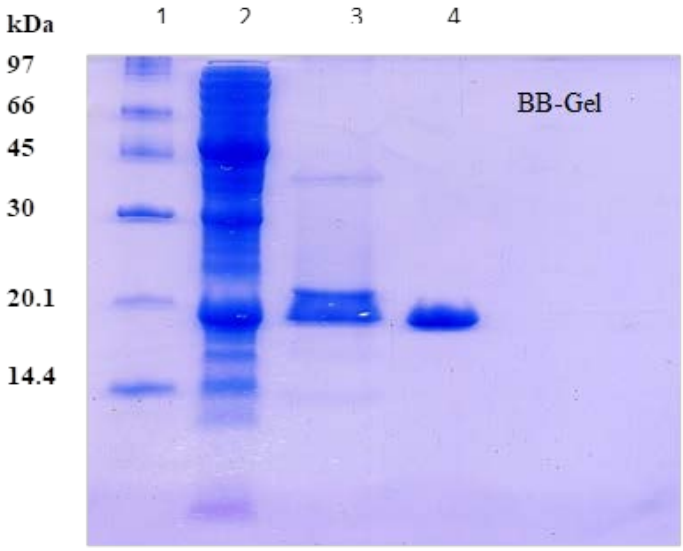


Figure 9B. SDS–PAGE results from peak fraction of Gel Filtration chromatography of BB subunits: lane 1 is the MW marker; lane 2 is periplasmic fraction; lane 3 is the GF fraction (19); and lane 4 is the GF f21.

Figure 10 below shows the results for BA subunits. Figure 10A shows the chromatogram of the gel filtration purification at 280 nm absorbance, and Figure 10B, on the next page, shows the SDS gel results of fractions from the Gel filtration purification step. GF Peak was located between fraction numbers 19 and 24.

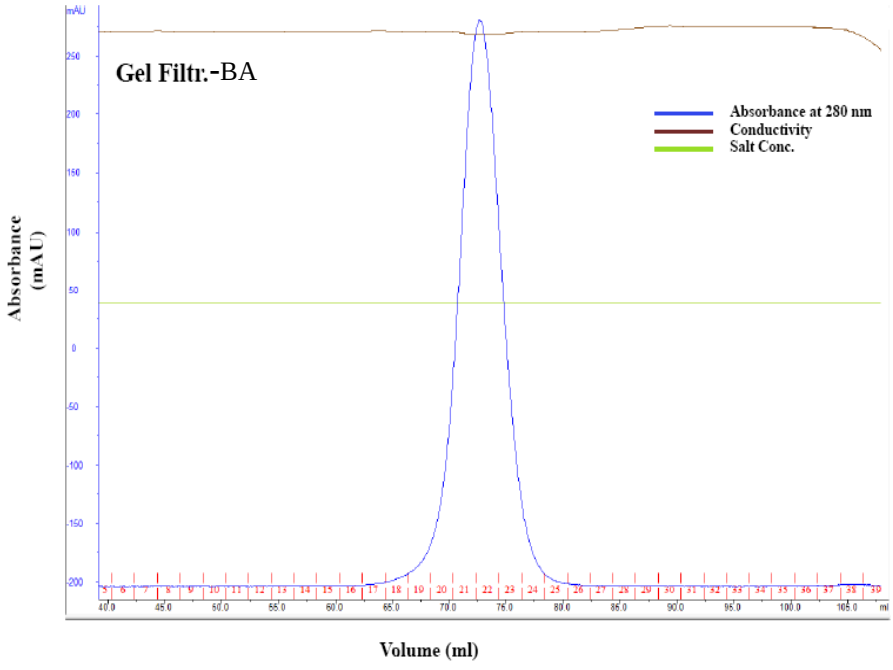


Figure 10A. Gel Filtration chromatography results for BA subunits: absorbance at 280 nm and peak between fraction numbers 19 and 24.

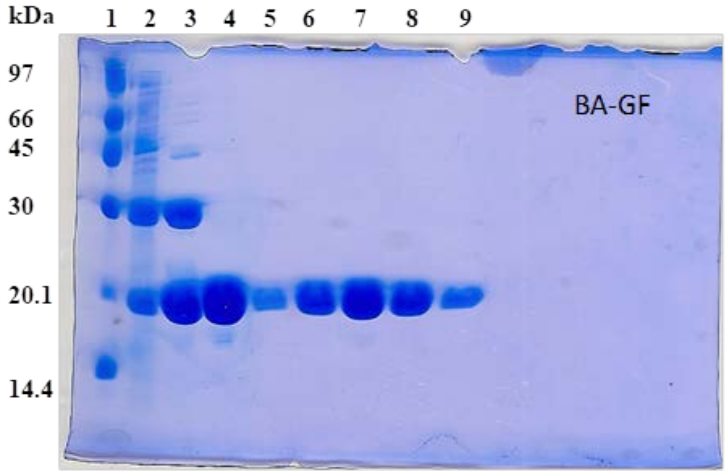


Figure 10B. SDS–PAGE results from peak fraction of Gel Filtration chromatography of BA subunits: lane1 is the MW marker; lane 2 is periplasmic fraction; lane 3 is S30Q sample; lane 4 is MQ sample; lane 5 is the GF fraction (19); lane 6 is the GF fraction (20); lane 7 is the GF fraction (21); lane 8 is GF fraction (22); and lane 9 is the GF fraction 24.

3.6 Results of Circular Dichroism (CD) Spectroscopy

Both chemical and thermal stability tests were done for the four subunits using (GdnHCl) and heat treatment between 30 °C to 90 °C, respectively. Results of the chemical and thermal denaturing of the four subunits were shown in Figure 11 and Figure 12 respectively. Figure 13 shows the fraction percentage of folded and unfolded protein states.

3.6.1 Chemical denaturation of the CssA_dscA subunits.

Figure 11 and Table 8 on the next page show the chemical denaturation of the four subunits. The AA subunits denatured or unfolded around 1M of GdnHCl, while the subunit AB unfolded around 2.25 M of GdnHCl. On the other hand, the BB subunit was unfolded around 3 M of GdnHCl. Finally, the BA subunit was stable for the highest concentration of 4 M Guanidine Hydrochloride. Note that hetero-polymerization of any of the two subunits, AB and BA gives higher chemical stability against the GdnHCl denaturation compared to the homo-polymerization, AA or BB. These results favour the presence of hetero-polymerization of the subunits, which would be used by the ETEC bacteria to form their effective structures.

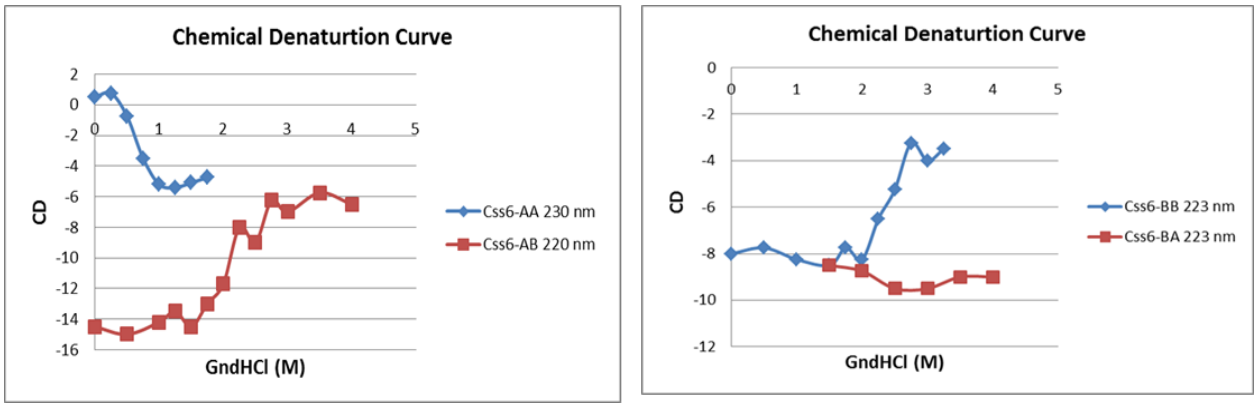


Figure 11. Circular Dichroism (CD) results of chemical denaturing of the four subunits.

Table 8. *Summary of Chemical stability using (GdnHCl) needed for protein denaturation*

Subunit	GdnHCl (M)
Css6-BA	4
Css6-BB	3
Css6-AB	2.25
Css6-AA	1

3.6.2 Thermal stability

Figure 12 in the next page, shows the results of the studied protein subunits which were heated from 30 °C to 90 °C in the absence of GdnHCl. Results of thermal denaturation showed that the AA subunits started to denature at 53°C, AB started at 75°C, BB started at 58°C, and finally, BA started at 71 °C. Here it is noted that hetero-polymerization of any two subunits, AB and BA gave higher stability against thermal denaturation compared to the homo-polymerization, AA or BB. These results favour the presence of hetero-polymerization of these subunits as the effective structures to be used by the ETEC Bacteria. The summary of melting temperature needed for protein denaturation is presented in Table 9 below.

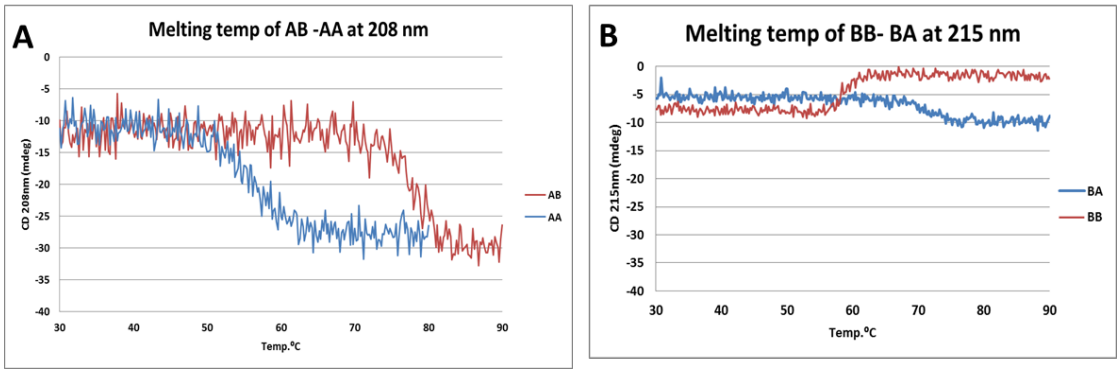


Figure 12. CD results of four subunits**.** Melting denaturing of AB-AA subunits at 208 nm showed in part A at left side. Melting denaturation of BB-BA subunits at 215 nm showed in part B at right side.

Table 9. *Summary of Melting Temperature needed for protein denaturation.*

Subunit	Temp (°C)
Css6-BA	71°
Css6-AB	75°
Css6-BB	58°
Css6-AA	53°

3.6.3 Calculated free energy change, ΔG (kJ/mol)

Figure 13 on the following page**,** shows the ellipticity curves plotted over a range of temperatures (K) and normalized between 0-1; representing the fraction (%) of protein in folded and unfolded states. The calculated free energy change **ΔG** (kJ/mol) is provided at the melting temperature (Tm), and is shown in Table 9 above, for each of the four subunits.

The free energy change is calculated at the melting temperature for the four subunits. ΔG was for AA, 0.22 (kJ/mol); for AB, the ΔG was 0.570 (kJ/mol). For BB, ΔG was 0.08 (kJ/mol); while for BA, the ΔG was 2.24 (kJ/mol). Table 10 on the next page summarises and compares the ΔG results based on thermal denaturation curves. Here it is noted that BA was the most stable subunit, then AB followed by BB and finally AA subunits. This suggests that hetero-polymerized subunits are more heat stable than homo-polymerized ones.

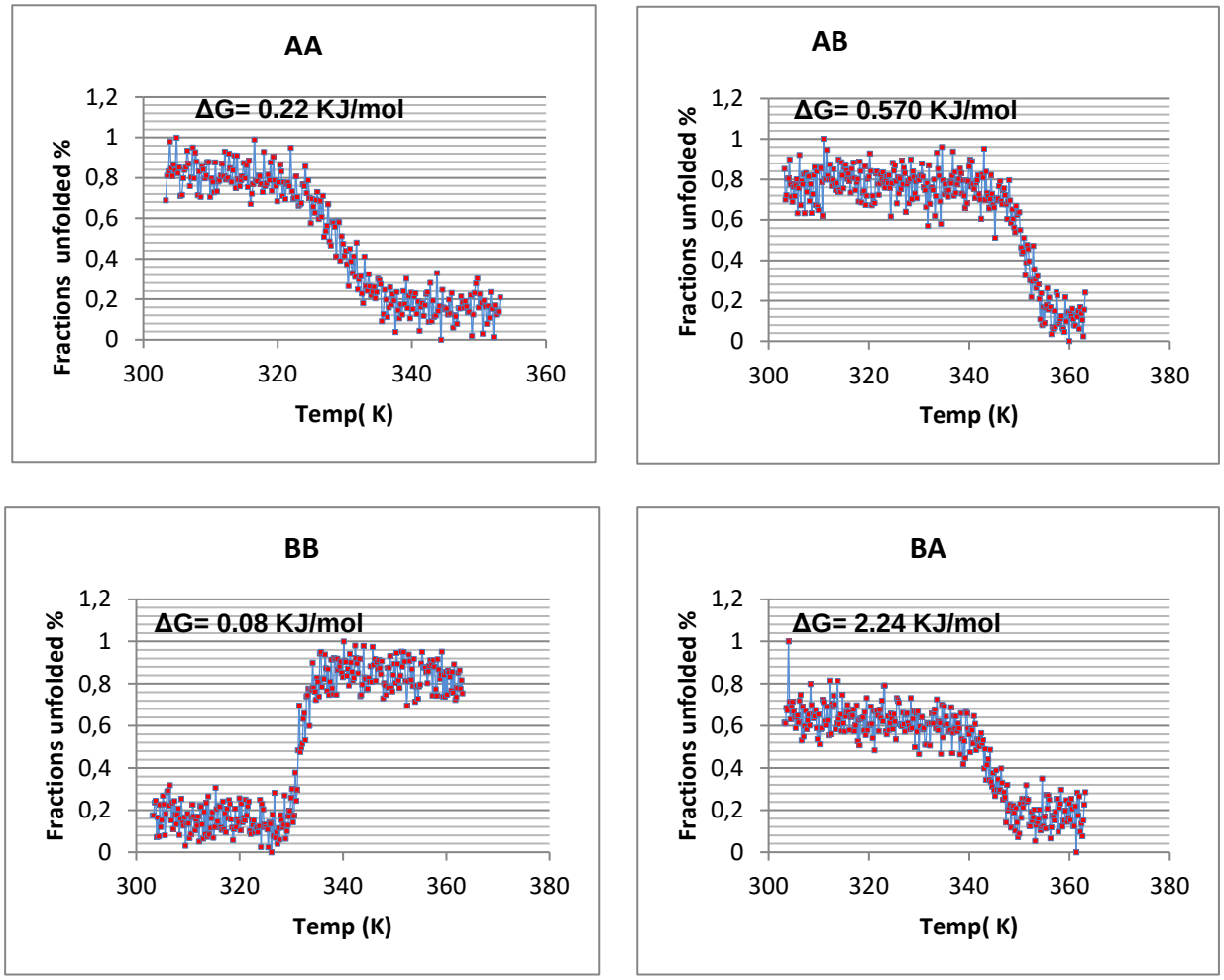


Figure 13. Fractions for folded and unfolded percentage of the four subunits, AA, AB, BB and BA as function of Temperature (K). Calculated free energy for unfolding at melting temperature is provided.

Table 10. *Summary free energy change, ΔG (KJ/mol), needed for thermal denaturation at melting point (Tm) of each subunit.*

Subunit	ΔG(KJ)
Css6- BA	**2.24**
Css6- AB	**0.570**
Css6 -BB	**0.080**
Css6-AA	**0.220**

4 Conclusions

The Donor strand complementation (DSC) theory was successfully implemented to study the four candidate fimbrial subunits based on previous project work. These four fimbrial subunits (AA, AB, BB and BA) were constructed using PCR, then protein expressed and purified using liquid chromatography techniques, after that, they have been studied chemically and thermally in a successful manner.

The protein expressions of the hetero-polymerized subunits (AB and BA) were stronger than the homo-polymerized subunits (AA and BB). Previous studies indicated that the expression of subunits (AA) and (BB) is in equal amounts but the latter (BB) is more immunogenic than (AA) (Ghosal, Bhowmick et al. 2009). These last finding regarding the equal expression of both subunits could strengthen our results about the predicted functional unit, favouring model 3 (M3) in our hypothesis as shown in Figure 2 previously. This prediction is favouring the formation of the hetero-polymerization of alternating both subunit A and B together**.** The other results of chemical denaturation, thermal stability and free energy change that were summarized in Tables, 8, 9 and 10, respectively, are all generally strengthening the possibility of the hetero–polymerization of alternating both subunits to be more stable than the homo-polymerization of AA or BB.

Further studies are needed to get better understanding of the mechanism of ETEC attachment to the host intestinal cells. These studies include crystallization, solving the 3D structure of the crystals, and studies on a candidate DNA vaccination strategy. In addition, other studies dealing with potentials treatment methods for inhibiting the formation of these colonization factors are needed.

5 References

ABU-ELYAZEED, R., WIERZBA, T. F., MOURAD, A. S., PERUSKI, L. F., KAY, B. A., RAO, M., CHURILLA, A. M., BOURGEOIS, A. L., MORTAGY, A. K., KAMAL, S. M., SAVARINO, S. J., CAMPBELL, J. R., MURPHY, J. R., NAFICY, A. & CLEMENS, J. D. 1999. Epidemiology of enterotoxigenic Escherichia coli diarrhea in a pediatric cohort in a periurban area of lower Egypt. J Infect Dis, 179, 382-9.

ALBERT, M. J., QADRI, F., WAHED, M. A., AHMED, T., RAHMAN, A. S., AHMED, F., BHUIYAN, N. A., ZAMAN, K., BAQUI, A. H., CLEMENS, J. D. & BLACK, R. E. 2003. Supplementation with zinc, but not vitamin A, improves seroconversion to vibriocidal antibody in children given an oral cholera vaccine. J Infect Dis, 187, 909-13.

BUSCH, A. & WAKSMAN, G. 2012. Chaperone-usher pathways: diversity and pilus assembly mechanism. Philos Trans R Soc Lond B Biol Sci, 367, 1112-22.

BYRD, W. & CASSELS, F. J. 2003. Mucosal immunization of BALB/c mice using enterotoxigenic Escherichia coli colonization factors CFA/I and CS6 administered with and without a mutant heat-labile enterotoxin. Vaccine, 21, 1884-93.

COHEN, D., TOBIAS, J., SPUNGIN-BIALIK, A., SELA, T., KAYOUF, R., VOLOVIK, Y. YAVZORI, M., EPHROS, M. 2010. Phenotypic characteristics of enterotoxigenic Escherichia coli associated with acute diarrhea among Israeli young adults. Foodborne Pathogens and Disease 7, 1159-1164

Creighton, T. (1997). Protein structure. 2nd ed. Oxford: IRL Press at Oxford University Press

GHOSAL, A., BHOWMICK, R., BANERJEE, R., GANGULY, S., YAMASAKI, S., RAMAMURTHY, T., HAMABATA, T. & CHATTERJEE, N. S. 2009. Characterization and studies of the cellular interaction of native colonization factor CS6 purified from a clinical isolate of enterotoxigenic Escherichia coli. Infect Immun, 77, 2125-35.

GIUGLIANI, E. R. & VICTORIA, C. G. 2000. [Complementary feeding]. J Pediatr (Rio J), 76 Suppl 3, S253-62.

GUERENA-BURGUENO, F., HALL, E. R., TAYLOR, D. N., CASSELS, F. J., SCOTT, D. A., WOLF, M. K., ROBERTS, Z. J., NESTEROVA, G. V., ALVING, C. R. & GLENN, G. M. 2002. Safety and immunogenicity of a prototype enterotoxigenic Escherichia coli vaccine administered transcutaneously. Infect Immun, 70, 1874-80.

GUPTA, S. K., KECK, J., RAM, P. K., CRUMP, J. A., MILLER, M. A. & MINTZ, E. D. 2008. Part III. Analysis of data gaps pertaining to enterotoxigenic Escherichia coli infections in low and medium human development index countries, 1984-2005. Epidemiol Infect, 136, 721-38.

JANSSON, L., TOBIAS, J., JAREFJALL, C., LEBENS, M., SVENNERHOLM, A. M. & TENEBERG, S. 2009. Sulfatide recognition by colonization factor antigen CS6 from enterotoxigenic Escherichia coli. PLoS One, 4, e4487.

JIANG, Z. D., LOWE, B., VERENKAR, M. P., ASHLEY, D., STEFFEN, R., TORNIEPORTH, N., VON SONNENBURG, F., WAIYAKI, P. & DUPONT, H. L. 2002. Prevalence of enteric pathogens among international travelers with diarrhea acquired in Kenya (Mombasa), India (Goa), or Jamaica (Montego Bay). J Infect Dis, 185, 497-502.

KAPER, J. B., NATARO, J. P. & MOBLEY, H. L. 2004. Pathogenic Escherichia coli. Nat Rev Microbiol, 2, 123-40.

Lee, D., Hong, Y., Choe, E., Lee, S., Kim, S., Lee, H., Oh, J., Shin, H. and Pyun, Y. (2005). A thermodynamic study of mesophilic, thermophilic, and hyperthermophilicl-arabinose isomerases: The effects of divalent metal ions on protein stability at elevated temperatures. FEBS Letters, 579(5), pp.1261-1266

LONG, K. Z., WOOD, J. W., VASQUEZ GARIBY, E., WEISS, K. M., MATHEWSON, J. J., DE LA CABADA, F. J., DUPONT, H. L. & WILSON, R. A. 1994. Proportional hazards analysis of diarrhea due to enterotoxigenic Escherichia coli and breast feeding in a cohort of urban Mexican children. Am J Epidemiol, 139, 193-205.

NICKLASSON, M., SJOLING, A., LEBENS, M., TOBIAS, J., JANZON, A., BRIVE, L. & SVENNERHOLM, A. M. 2008. Mutations in the periplasmic chaperone leading to loss of surface expression of the colonization factor CS6 in enterotoxigenic Escherichia coli (ETEC) clinical isolates. Microb Pathog, 44, 246-54.

QADRI, F., SAHA, A., AHMED, T., AL TARIQUE, A., BEGUM, Y. A. & SVENNERHOLM, A. M. 2007. Disease burden due to enterotoxigenic Escherichia coli in the first 2 years of life in an urban community in Bangladesh. Infect Immun, 75, 3961-8.

QADRI, F., SVENNERHOLM, A. M., FARUQUE, A. S. & SACK, R. B. 2005. Enterotoxigenic Escherichia coli in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. Clin Microbiol Rev, 18, 465-83.

RAO, M. R., WIERZBA, T. F., SAVARINO, S. J., ABU-ELYAZEED, R., EL-GHOREB, N., HALL, E. R., NAFICY, A., ABDEL-MESSIH, I., FRENCK, R. W., JR., SVENNERHOLM, A. M. & CLEMENS, J. D. 2005. Serologic correlates of protection against enterotoxigenic Escherichia coli diarrhea. J Infect Dis, 191, 562-70.

REMAUT, H. & WAKSMAN, G. 2004. Structural biology of bacterial pathogenesis. Curr Opin Struct Biol, 14, 161-70.

RIVERA, F. P., OCHOA, T. J., MAVES, R. C., BERNAL, M., MEDINA, A. M., MEZA, R., BARLETTA, F., MERCADO, E., ECKER, L., GIL, A. I., HALL, E. R., HUICHO, L. & LANATA, C. F. 2010. Genotypic and phenotypic characterization of enterotoxigenic Escherichia coli strains isolated from Peruvian children. J Clin Microbiol, 48, 3198-203.

SABUI, S., GHOSAL, A., DUTTA, S., GHOSH, A., RAMAMURTHY, T., NATARO, J. P., HAMABATA, T. & CHATTERJEE, N. S. 2010. Allelic variation in colonization factor CS6 of enterotoxigenic Escherichia coli isolated from patients with acute diarrhoea and controls. J Med Microbiol, 59, 770-9.

SEARS, C. L. & KAPER, J. B. 1996. Enteric bacterial toxins: mechanisms of action and linkage to intestinal secretion. Microbiol Rev, 60, 167-215.

SHAHEEN, H. I., KAMAL, K. A., WASFY, M. O., EL-GHORAB, N. M., LOWE, B., STEFFEN, R., KODKANI, N., AMSLER, L., WAIYAKI, P., DAVID, J. C., KHALIL, S. B. & PERUSKI, L. F., JR. 2003. Phenotypic diversity of enterotoxigenic Escherichia coli (ETEC) isolated from cases of travelers' diarrhea in Kenya. Int J Infect Dis, 7, 35-8.

TOBIAS, J., HOLMGREN, J., HELLMAN, M., NYGREN, E., LEBENS, M. & SVENNERHOLM, A. M. 2010. Over-expression of major colonization factors of enterotoxigenic Escherichia coli, alone or together, on non-toxigenic E. coli bacteria. Vaccine, 28, 6977-84.

TOBIAS, J., LEBENS, M., KALLGARD, S., NICKLASSON, M. & SVENNERHOLM, A. M. 2008. Role of different genes in the CS6 operon for surface expression of Enterotoxigenic Escherichia coli colonization factor CS6. Vaccine, 26, 5373-80.

WAKSMAN, G. & HULTGREN, S. J. 2009. Structural biology of the chaperone-usher pathway of pilus biogenesis. Nat Rev Microbiol, 7, 765-74.

YATES, J. 2005. Traveler's diarrhea. Am Fam Physician, 71, 2095-100.

ZAVIALOV, A. V. & KNIGHT, S. D. 2007. A novel self-capping mechanism controls aggregation of periplasmic chaperone Caf1M. Mol Microbiol, 64, 153-64.

ZAVIALOV, A. V., BATCHIKOVA, N. V., KORPELA, T., PETROVSKAYA, L. E., KOROBKO, V. G., KERSLEY, J., MACINTYRE, S. & ZAV'YALOV, V. P. 2001. Secretion of recombinant proteins via the chaperone/usher pathway in Escherichia coli. Appl Environ Microbiol, 67, 1805-14.

ZAVIALOV, A. V., KERSLEY, J., KORPELA, T., ZAV'YALOV, V. P., MACINTYRE, S. & KNIGHT, S. D. 2002. Donor strand complementation mechanism in the biogenesis of non-pilus systems. Mol Microbiol, 45, 983-95.

ZAVIALOV, A., ZAV'YALOVA, G., KORPELA, T. & ZAV'YALOV, V. 2007. FGL chaperone-assembled fimbrial polyadhesins: anti-immune armament of Gram-negative bacterial pathogens. FEMS Microbiol Rev, 31, 478-514.

6 Popular science summary

Enterotoxigenic *Escherichia coli* (ETEC) are one of common pathogens causing diarrhoea in people worldwide. This bacterium is affecting mainly young children under 5 years of old in developing countries, which can lead to their death and be regarded as the second most common cause of diarrhoea after retrovirus in children. For travellers, it causes the famous disease “traveller’s diarrhoea”. Adult visitors and tourists from North America and Europe to endemic areas of developing countries are often affected with traveller’s diarrhoea caused mainly by ETEC and carry back the pathogen to their home lands.

Regions of rapid population growth especially in suburban regions with poor quality water, poor infrastructures facilities and lack of sewage systems will increase the risk factors of ETEC infections. The main source for infection with ETEC is the faecal-oral route from contaminated food, water and beverages due to poor hygienic conditions.

Clinical Feature

Most clinical cases of diarrhoea caused by ETEC can occur early after eight hours from the infection and last about four days without treatment (self-limited), while symptoms vary from mild to severe. The pathophysiology of the ETEC infection resembles the disease caused by Vibrio cholera (causes cholera disease) and the symptoms are usually identical especially in adult patients.

Bacteriology and virulence factors

Enterotoxigenic *Escherichia coli* (ETEC) belong to the Enterobacteriaceae family, which is mostly Gram-negative, non-spore-forming, motile, lactose fermentative rod shaped organisms.

In order to induce the disease, the ETEC strains should attach or colonize the small intestine mucosa after being ingested by individuals. This attachment to the host enteric epithelial cells is facilitated by the presence of surface antigens called colonization factor antigens (CFAs) or coli surface antigens (CSs). These factors are mostly immunogenic proteins that bind to receptors on the host intestinal mucosa.

The frequently isolated ETEC strains from patients have an immunogenic surface protein colonization factor CS6. This surface antigen CS6 is helping in adhesion to the host small intestinal epithelium enabling ETEC to produces their two main toxins, heat-labile (LT) and/ or heat-stable (ST) enterotoxin.

Aim of This Study

This study was conducted as the first step to understand the possible basic structures of the CS6 gene, which responsible mainly in attachment of ETEC to the patient intestinal cells. Revealing the structure and assembly of this CS6 gene will help in understanding the pathogenesis, diagnosis and future treatment methods. Vaccination as well can be studied and developed as a preventive method against ETEC.

7 Acknowledgements

Firstly and Finally Praise be to Allah, Lord of the Worlds for all the times of thesis work, Interruption Period, accomplishing, failure and success in this work and all aspect of life.

I would like to express my sincere gratitude and appreciation to my supervisor Prof. Anton Zavialov, for his inspiration, knowledge and patience during my work with his fantastic research and group. Same thanks for Prof. Jerry Ståhlberg for his help and support during revising and examination of my thesis, from Department of Molecular Sciences at SLU, Uppsala.

Special Thanks to Professor Dr Volkmar Passoth, Department of Molecular Sciences, SLU, Uppsala for his assistance and help to accomplish this thesis after Interruption time. Great thanks to the Department of Molecular Sciences at SLU Uppsala, for applying resources, equipment and their cooperation to apply this work.

Very special Thanks to Husam Babikir, for his great friendship and times we spent together during master studies and for the endless talks about everything between earth and sky.

A special thanks from the bottom of my heart for all times to the soul of my father, *may Allah bless his soul,* and to my struggling mother for their continuous and incredible support throughout my life, and for all my brothers and sisters, Specially Mohammad and Mustafa for their support and encourage.

Finally and always, for my lover, friend and soul mate, my beloved wife, Enas, a special thanks from the bottom of my heart, but for you my heart has no bottom, and for my small angels and part of my heart, Tala and Qais with all my love.